

RA Marsh V.L., Martin S.L., McConnachie L.J., McLeay K., McMurray A.A.,
RA Milne S., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S.,
RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
RA Swann M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA Rogers J.;
RL "The DNA sequence and comparative analysis of human chromosome 20.";
RL Nature 414:865-871(2001).
CC -!- FUNCTION: MAY PLAY A ROLE IN FATTY ACID OXIDATION RATHER THAN
CC FORMATION OF FATTY ACIDS. MAY MEDATE NEF-INDUCED DOWN-REGULATION
CC OF CD4.
CC -!- CATALYTIC ACTIVITY: Palmitoyl-CoA + H(2)O -> CoA + palmitate.
CC -!- SUBUNIT: INTERACTS WITH HIV-1 NEF.
CC -!- SUBCELLULAR LOCATION: Peroxisomal.
CC -!- SIMILARITY: BELONGS TO THE C/M/P THIOESTER HYDROLASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL AF014404; AAB71665.1; -;
CC EMBL X86032; CAA60024.1; -;
CC EMBL AF124264; AAD27616.1; -;
CC EMBL AL008726; CAA1502.1; -;
CC HSP: P23911; 1C8U.
CC InterPro: IPR003703; Acyl-CoA thio.
CC Pfam: PF05551; Acyl-CoA thio; 1.
CC Hydrolase; Serine esterase; Peroxisome.
CC ACT_SITE 78 78 BY SIMILARITY.
CC SITE 317 319 MICROBODY TARGETING SIGNAL (POTENTIAL).
CC CONFLICT 291 293 LAR -> VMS (IN REF. 2).
CC CONFLICT 319 319 L -> R (IN REF. 2).
CC SEQUENCE 319 AA; 35914 MW; 8345C6E5EABF3326 CRC64;

Query Match 97.3%; Score 1621; DB 1; Length 319;
Best Local Similarity 100.0%; Pred. No. 2.1e-138;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
11 PFGDLRSVLVTTLNLEPLDEDLFRGRHYWVPAKRLFGGQIVGQALVAAKSVSDVHVH 70
19 PFGDLRSVLVTTLNLEPLDEDLFRGRHYWVPAKRLFGGQIVGQALVAAKSVSDVHVH 78
71 SLHCYFVRAGDPKPLVLYOVERTRTGSSFSVRSYKAVQHGKPIFCQASFOQAQSPMOH 130
79 SLHCYFVRAGDPKPLVLYOVERTRTGSSFSVRSYKAVQHGKPIFCQASFOQAQSPMOH 138
131 QFSMPTVPPEELLDCETLIDQYLRDPNLQKRYPLALNRIAAQEVPIEKVPNPISQL 190
139 QFSMPTVPPEELLDCETLIDQYLRDPNLQKRYPLALNRIAAQEVPIEKVPNPISQL 198
191 QRMPEKQMFVWVARGYIGSGDMKHCVCVAYISDYAFLCTALLPHQWQKRVHVMVSLDHS 250
199 QRMPEKQMFVWVARGYIGSGDMKHCVCVAYISDYAFLCTALLPHQWQKRVHVMVSLDHS 258
251 MNFHPAPFRADHMLYECESFWAGGSGRLVHGLRWDRDGLVLTCAQGVIRKVPQVSEK 310
259 MNFHPAPFRADHMLYECESFWAGGSGRLVHGLRWDRDGLVLTCAQGVIRKVPQVSEK 318
311 L 311
319 L 319

ALIGNMENTS

RESULT 1
ID PTE1_HUMAN STANDARD; PRT; 319 AA.
16-OCT-2001 (Rel. 40, Created)
16-OCT-2001 (Rel. 40, Last sequence update)
01-MAR-2002 (Rel. 41, Last annotation update)
DE Peroxisomal acyl-coenzyme A thioester hydrolase 1 (EC 3.1.1.2.2)
DE (Peroxisomal long-chain acyl-CoA thioesterase 1) (HIV-Nef associated
DE acyl CoA thioesterase) (thioesterase II) (HTE).
GN PTE1 OR HNAACE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97445158; PubMed=9299485;
RA Watanabe H., Shiratori T., Shoji H., Miyatake S., Okazaki Y.,
RA Ikuta K., Sato T., Saito T.;
RT "A novel acyl-CoA thioesterase enhances its enzymatic activity by
RT direct binding with HIV Nef.";
RL Biochem. Biophys. Res. Commun. 238:234-239(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymphoid;
RX MEDLINE=97298085; PubMed=10092594;
RA Liu L.X., Margottin F., LeGall S., Schwartz O., Selig L., Benarous R.,
RA Benichou S.;
RT "Binding of HIV-1 Nef to a novel thioesterase enzyme correlates with
RT Nef-mediated CD4 down-regulation.";
RL J. Biol. Chem. 272:13779-13785(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RX MEDLINE=99194760; PubMed=10092594;
RA Jones J.M., Nau K., Geraghty M.T., Erdmann R., Gould S.J.;
RT "Identification of peroxisomal acyl-CoA thioesterases in yeast and
RT humans.";
RL J. Biol. Chem. 274:9216-9223(1999).
RN [4]
RP SEQUENCE FROM N.A.
RA DeLoukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA Buck D., Burrill W., Butler A.P., Carder C., Carter N.P.,
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA Clegg S., Cobley V.E., Collier R.E., Connor R., Corby N.R.,
RA Coulson A., Coville G.J., Deadman R., Dhali P., Dunn M.,
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA Graffham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA LeHvasalaho M.H., Leversha M., Lloyd C., Lloyd D.M., Lovell J.D.,



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 19, 2002, 13:18:34 ; Search time 17.44 Seconds
(without alignments)
1713.520 Million cell updates/sec

Title: US-09-766-366-1
Perfect score: 1666
Sequence: 1 MGRAVATAALPPGDLRSVLV.....VTCAQEGVIRVQPVSESKL 311

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues
T number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 71:.*
1: PIR1:.*
2: PIR2:.*
3: PIR3:.*
4: PIR4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1621	97.3	319	1 JC5644	acyl-CoA thioleste
2	612	36.7	300	2 AI3272	acyl-CoA thioester
3	583.5	35.0	294	2 AC2915	acyl-CoA thioester
4	583.5	35.0	316	2 F97689	acyl-CoA thioester
5	568	34.1	289	2 B83153	acyl-CoA thioester
6	567	34.0	343	2 T29819	hypothetical prote
7	559	33.6	310	2 A82734	conserved hypothet
8	558	33.5	286	2 AE0381	acyl-CoA thioester
9	553	33.2	286	1 D64775	acyl-CoA thioleste
10	553	33.2	286	2 B90692	acyl-CoA thioester
11	553	33.2	286	2 F85542	acyl-CoA thioester
12	543	32.6	286	2 A82248	acyl-CoA thioester
13	536	32.2	286	2 AD0560	acyl-CoA thioester
14	518.5	31.1	430	2 T25623	hypothetical prote
15	501.5	30.1	284	2 G87701	acyl-CoA thioester
16	499.5	30.0	320	2 C86148	hypothetical prote
17	460	27.6	286	1 B64047	acyl-CoA thioleste
18	442.5	26.6	281	1 D70370	probable acyl-CoA
19	440.5	26.4	465	2 T15540	hypothetical prote
20	327.5	19.7	365	2 T27460	hypothetical prote
21	322	19.3	213	2 T01216	hypothetical prote
22	311.5	18.7	297	2 H87068	acyl CoA thioester
23	294.5	17.7	300	1 H70557	probable acyl-CoA
24	273	16.4	349	1 S52763	hypothetical prote
25	243.5	14.6	276	2 T15536	hypothetical prote
26	115.5	6.9	323	2 D87472	hypothetical prote
27	112.5	6.8	250	2 T15887	hypothetical prote
28	101	6.1	265	2 D83287	hypothetical prote
29	100.5	6.0	908	2 T25035	hypothetical prote

30	99.5	6.0	417	1 WMBEHA	52K immediate-earl
31	93	5.6	354	2 T10371	hypothetical prote
32	93	5.6	354	2 C34526	ORF3 protein - Org
33	92	5.5	276	2 S57656	glutenin low molec
34	91.5	5.5	1070	2 JC4593	protein-tyrosine k
35	91	5.5	744	2 AC1970	hypothetical prote
36	90.5	5.4	1234	2 T00363	hypothetical prote
37	90.5	5.4	2133	2 T30637	hypothetical prote
38	90	5.4	454	2 A69017	dihydroorotase - M
39	89.5	5.4	1502	1 R6BYH1	CYC1/CYP3 transcri
40	89	5.3	1279	1 B69681	peptide synthetase
41	89	5.3	1978	2 T07081	acetyl-CoA carboxy
42	89	5.3	2304	2 T07920	probable acetyl-Co
43	89	5.3	2325	2 T07235	acetyl-CoA carboxy
44	88.5	5.3	644	2 F82145	conserved hypothet
45	88	5.3	499	2 F86206	hypothetical prote

ALIGNMENTS

RESULT 1
JC5644
acyl-CoA thiolesterase (EC 3.1.2.-) III, peroxisomal - human
N:Alternate names: HIV Nef-binding protein
N:Contains: palmitoyl-CoA hydrolase (EC 3.1.2.2)
C:Species: Homo sapiens (man)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 05-May-2000
C:Accession: JC5644
R:Watanabe, H.; Shiratori, T.; Shoji, H.; Miyatake, S.; Okazaki, Y.; Ikuta, K.; Sata
Biochem. Biophys. Res. Commun. 238, 234-239, 1997
A:Title: A novel acyl-CoA thioesterase enhances its enzymatic activity by direct bl
A:Reference number: JC5644; MUID:97445158
A:Accession: JC5644
A:Molecule type: mRNA
A:Residues: 1-319 <WAT>
A:Cross-references: DDBJ:AF014404; NID:G2318124; PIDN:ABE71665.1; PID:G2318125
C:Comment: This enzyme regulates membrane sorting and subcellular morphology.
C:Genetics:
A:Gene: GDB:PTE1
A:Cross-references: GDB:9954875
C:Superfamily: acyl-CoA thiolesterase II
C:Keywords: coenzyme A; peroxisome; thiolester hydrolase
F:317-319/Region: peroxisome/glyoxysome location signal (S-[RKH]-L) motif
F:78/Active site: His #status predicted

Query Match	97.3%	Score 1621;	DB 1;	Length 319;
Best Local Similarity	100.0%;	Pred. No. 3.1e-136;		
Matches 301;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	11	PPGDLRSVLVTVTLNLEPLDEDLFRGRHYWVPKRLFGGQIVGQALVAAAKSVSEVDVHVH	70	
Db	19	PPGDLRSVLVTVTLNLEPLDEDLFRGRHYWVPKRLFGGQIVGQALVAAAKSVSEVDVHVH	78	
QY	71	SLHCYFVRAGDPKLPVLYQVTRTGTSSFSVRSVRAVHGKPIFTICQASFQQAQSPMQH	130	
Db	79	SLHCYFVRAGDPKLPVLYQVTRTGTSSFSVRSVRAVHGKPIFTICQASFQQAQSPMQH	138	
QY	131	QFSMPTVPPPEELDCETFLIDQYLRDPLNOKRYPLALNRFAAQEVPDIEIKPVNPSPLSQL	190	
Db	139	QFSMPTVPPPEELDCETFLIDQYLRDPLNOKRYPLALNRFAAQEVPDIEIKPVNPSPLSQL	198	
QY	191	QRMEPKQMFWRARGYIGEGDMKHCVAAYISDYAFGLTALLPHQHQHVKHFMVSLDHS	250	
Db	199	QRMEPKQMFWRARGYIGEGDMKHCVAAYISDYAFGLTALLPHQHQHVKHFMVSLDHS	258	
QY	251	MWFHAPRADHWMLYECESPNAGSGRLVHGLRWODGVLAVTCAQEGVIRVQPVSESK	310	
Db	259	MWFHAPRADHWMLYECESPNAGSGRLVHGLRWODGVLAVTCAQEGVIRVQPVSESK	318	
QY	311	L 311		

Db 319 L 319

RESULT 2

AI3272

acyl-CoA thioesterase II (EC 3.1.2.-) [imported] - Brucella melitensis (strain 16M)

C:Species: Brucella melitensis

C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 15-Feb-2002

C:Accession: AI3272

R:DelVecchio, V.G.; Kapatal, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova, .; Mazur, M.; Goldsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002

A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis

A:Reference number: AB3252; PMID:11756688

A:Accession: AI3272

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-300 <KUR>

A:Cross-references: GB:AE008917; PIDN:AAL51348.1; PID:g17982047; GSPDB:GN00190

A:Experimental source: strain 16M

C:Genetics:

A:Gene: BMEI0166

A:Map position: 1

C:Superfamily: acyl-CoA thioesterase II

C:Keywords: thiolester hydrolase

Query Match 36.7%; Score 612; DB 2; Length 300;

Best Local Similarity 44.4%; Pred. No. 1.1e-46;

Matches 135; Conservative 52; Mismatches 89; Indels 28; Gaps 9;

Qy 7 TAALPGDLRSVLTVLNLEPLDEDLFRGRHWYVPAKRLFGGQIVGOALVAAAKSVSED 66

Db 12 TAA-----MRELL--SILDELTLEMDLFRGNSQVQVFGGQIVGOALIAAQTVDPE 64

Qy 67 VVHSLHCFYFRAGDPKLPVLYQVTRTSGSFVSRSKAVOHGKPIFCQASFOQAQPS 126

Db 65 RHVSHLAYFVRGDAPIIIEVDRIKDSFSTRLVAKOHGKAFTLSASF-QIDEG 123

Qy 127 PMQHFSMPT-VPPPELDCETLIDQYL--RDPNLQKRYPLALNRIAAEVPIEIKPVN 183

Db 124 GLDHQMPPEGLPQPEQLVGDHDIKEKYLDMAPPGVKRYW-----ERERPIEIKPVS 175

Qy 184 PSPLSOLQRMPEKQWFWARCYIGGDMKHCVCVAAVYISDYAFGLTALLPHQVHF 243

Db 176 LTHYFSREKLEPVQVHWVVRAGLVDP-DRALQAAILAYLSMDTLTSLPHP---GRFIF 231

244 -----MVSLDHSWFFHAPRADHMLYECESPWAGSGRLVHGLWRQDGLVAVTCAQEG 298

232 DRDMQVASILDHAMWFRPCRLDMLLYTODAPASGARGFNRGALYTRDGLVLTASVAQEG 291

Qy 299 VIRV 302

Db 292 LIRV 295

RESULT 3

AC2915

acyl-CoA thioesterase II [imported] - Agrobacterium tumefaciens (strain C58, Dupont)

C:Species: Agrobacterium tumefaciens

C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 01-Feb-2002

C:Accession: AC2915

R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClellan ; Karp, P.; Romero, P.; Zhang, S.

A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A:Reference number: AB2577; PMID:11743193

A:Accession: AC2915

A:Status: preliminary

A:Molecule type: DNA

Query Match 35.0%; Score 583.5; DB 2; Length 316;

Best Local Similarity 42.3%; Pred. No. 3.9e-44;

Matches 121; Conservative 56; Mismatches 96; Indels 13; Gaps 6;

Qy 19 LVTTVLNLEPLDEDLFRGRHWYVPAKRLFGGQIVGOALVAAAKSVSEDVHSHLHCYFVR 78

Db 36 LIAT-LDLEKLEENLYRGSPQIGWQVFGQVIAQAALIAAQTVDVDDRFVHSLHAYFMR 94

Qy 79 AGDKPLVLYQVTRTSGSFVSRSKAVOHGKPIFCQASFOQAQPSPMOHQFSMPTVP 138

Db 95 PGDLPLVLYQVTRTSGSFVSRSKAVOHGKPIFCQASFOQAQPSPMOHQFSMPTVP 153

Qy 139 PPEELDCETLIDQYLDRDNLQKRYPLALNRIAAEVPIEIKPVNPSPLSOLQRMPEKQ 198

Db 154 PPEKLMSEEQMAAFL-----AKAPASIRKYWSNKRPIEIRPVSLTHYISKEKLEPQD 207

Qy 199 FWVRARGYIGEGDMKHCVCVAAVYISDYAFGLTALLPHQ---WOHKVHFMVSLDHSWFFHA 255

A:Residues: 1-294 <KUR>

A:Cross-references: GB:AE008688; PIDN:AAL43737.1; PID:g17741270; GSPDB:GN00186

A:Experimental source: strain C58 (Dupont)

C:Genetics:

A:Gene: tesB

A:Map position: circular chromosome

C:Superfamily: acyl-CoA thioesterase II

Query Match 35.0%; Score 583.5; DB 2; Length 294;

Best Local Similarity 42.3%; Pred. No. 3.6e-44;

Matches 121; Conservative 56; Mismatches 96; Indels 13; Gaps 6;

Qy 19 LVTTVLNLEPLDEDLFRGRHWYVPAKRLFGGQIVGOALVAAAKSVSEDVHSHLHCYFVR 78

Db 14 LIAT-LDLEKLEENLYRGSPQIGWQVFGQVIAQAALIAAQTVDVDDRFVHSLHAYFMR 72

Qy 79 AGDKPLVLYQVTRTSGSFVSRSKAVOHGKPIFCQASFOQAQPSPMOHQFSMPTVP 138

Db 73 PGDLPLVLYQVTRTSGSFVSRSKAVOHGKPIFCQASFOQAQPSPMOHQFSMPTVP 131

Qy 139 PPEELDCETLIDQYLDRDNLQKRYPLALNRIAAEVPIEIKPVNPSPLSOLQRMPEKQ 198

Db 132 PPEKLMSEEQMAAFL-----AKAPASIRKYWSNKRPIEIRPVSLTHYISKEKLEPQD 185

Qy 199 FWVRARGYIGEGDMKHCVCVAAVYISDYAFGLTALLPHQ---WOHKVHFMVSLDHSWFFHA 255

Db 186 IWVRAYGEV-PADPRLQSAIAYLSMDTLTSLYAHGTTTFDPSIQ-VASLDHAMWFRH 243

Qy 256 PFRADHMLYECESPWAGSGRLVHGLWRQDGLVAVTCAQEGVIR 301

Db 244 PCRLDMLLYTODSPSARGALTRGNIFTRQGLVAVASVAQEGVIR 289

RESULT 4

F97689

acyl-CoA thioesterase II (PA3942) [imported] - Agrobacterium tumefaciens (strain C58)

C:Species: Agrobacterium tumefaciens

C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 11-Jan-2002

C:Accession: F97689

R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; G A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markel Science 294, 2323-2328, 2001

A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacteri

A:Reference number: A97359; PMID:11743194

A:Accession: F97689

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-316 <KUR>

A:Cross-references: GB:AE007869; PIDN:AAK88471.1; PID:g15157976; GSPDB:GN00169

C:Genetics:

A:Gene: AGR_C4997

A:Map position: circular chromosome

C:Superfamily: acyl-CoA thioesterase II

Query Match 35.0%; Score 583.5; DB 2; Length 316;

Best Local Similarity 42.3%; Pred. No. 3.9e-44;

Matches 121; Conservative 56; Mismatches 96; Indels 13; Gaps 6;

Qy 19 LVTTVLNLEPLDEDLFRGRHWYVPAKRLFGGQIVGOALVAAAKSVSEDVHSHLHCYFVR 78

Db 36 LIAT-LDLEKLEENLYRGSPQIGWQVFGQVIAQAALIAAQTVDVDDRFVHSLHAYFMR 94

Qy 79 AGDKPLVLYQVTRTSGSFVSRSKAVOHGKPIFCQASFOQAQPSPMOHQFSMPTVP 138

Db 95 PGDLPLVLYQVTRTSGSFVSRSKAVOHGKPIFCQASFOQAQPSPMOHQFSMPTVP 153

Qy 139 PPEELDCETLIDQYLDRDNLQKRYPLALNRIAAEVPIEIKPVNPSPLSOLQRMPEKQ 198

Db 154 PPEKLMSEEQMAAFL-----AKAPASIRKYWSNKRPIEIRPVSLTHYISKEKLEPQD 207

Qy 199 FWVRARGYIGEGDMKHCVCVAAVYISDYAFGLTALLPHQ---WOHKVHFMVSLDHSWFFHA 255

Query Match 33.6%; Score 559; DB 2; Length 310;
Best Local Similarity 39.5%; Pred. No. 5.8e-42;
Matches 117; Conservative 54; Mismatches 107; Indels

A:Accession: AD0560
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-286 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD08924.1; PID:gl16501736; GSPDB:GN00176
C:Genetics: STY0508
C:Superfamily: acyl-CoA thiolesterase II

Query Match 32.2%; Score 536; DB 2; Length 286;
Best Local Similarity 41.8%; Pred. No. 5.7e-40;
Matches 123; Conservative 49; Mismatches 88; Indels 34; Gaps 9;

QY 22 TVLNLEPLEDEDFRGRHYWVPAKRLFGGQIVGOALVAAAKSVSESVHSLHCYFVRAGD 81
DB 10 TLLNLEKIEGLFRGQSEDLGLRGVGGVGOALVAAAKETVPEARLVHSHFHYFLRPGD 69
QY 82 KPLVLYOVERTRTGSSFSVRSVKAQHGKPIFICQASFOQAQSPMOHQFSMTVPVPPPE 141
DB 70 SOKPIYDEVLDRGNSFARRVAALQNGKPIFYWTASFQAEFG-FERQKTMPTAVGPE 128
QY 142 ELLDCETLIDQ---YLDRPNLQKRYPLALNRIAAOEVPTEIRPV---NP-----SPLSQ 189
DB 129 G-LPSETEIAOHLAPLPPILXEP-----LCDRPLEIRVEFHNPLKGVHVSAPVRQ 179
QY 190 LQRMPEKQMFVRRARGYIGEGDMKHCVCVAAVYISDYAFGLTALLPH--QWQHKVHFVMSL 247
DB 180 V-----WIRANGTVPD-DIRVHOYLLGYASDLNLFPLVALOPHGIGFLEKGIQIATI 229
QY 248 DHSMFHFAPFRADHMLYECESFWAGSGRLVHGRVLRQDGLVAVTCAOEGVIR 301
DB 230 DHSMFHFAPFRNEWLLYSVESASSARGFVRGEFTQDGLVAVTCAOEGVIR 283

RESULT 14

T25623
Hypothetical protein C37H5.13 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 19-May-2000
C:Accession: T25623
R:Davidson, S.; Gillam, B.
Submitted to the EMBL Data Library, February 1997
A:Description: The sequence of C. elegans cosmid C37H5.
A:Reference number: T25058
A:Accession: T25623
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-430 <DAY>
A:Cross-references: EMBL:U08315; PIDN:AAB42376.1; GSPDB:GN000023; CESP:C37H5.13
A:Experimental source: strain Bristol N2; clone C37H5
C:Genetics:
A:Gene: CESP:C37H5.13
A:Map position: 5
A:Introns: 14/3; 30/3; 80/3; 132/2; 222/3; 283/3; 352/2
C:Superfamily: acyl-CoA thiolesterase II

Query Match 31.1%; Score 518.5; DB 2; Length 430;
Best Local Similarity 37.0%; Pred. No. 3.4e-38;
Matches 121; Conservative 53; Mismatches 116; Indels 37; Gaps 8;

QY 14 DLRSVLVTTLNLEPLEDEDFRGRHYWVPAKRL---FGGQIVGOALVAAAKSVSESVH 70
DB 111 DIRAGLIDTFLNLRVDRKMLARLLKGRNSLPVYGGVIGVGOALSAATATVEVGFVN 170
QY 71 SLHCYFVRAGDKPLVLYOVERTRTGSSFSVRSVKAQHGKPIFICQASFOQAQSPMOH 130
DB 171 SLHSYFVQSGNVVERPILYQVDRIRDGSKSCTRLVRLALQDGEALFTVQISFHRPEADSI 230
QY 131 QFSMTVPVPPPELDCETLIDQYLRDPLN-----OKRYPLALNRIRIAOEVPIEI 179
DB 231 QLPMEVFPAPDSLEDSLDTFERIKKNANIPPAALAMIGFKQKRIIPAFPRI-----FSF 284

QY 180 KPVNPSPLSQORMEP-----KQMEFWRRARGYIGEGDMKHCVCVAAVYISDYAPL 228
DB 285 RPVDIDSYLCLKDDHTAGHGHTDAYRSYVWIKANENIGD-DPRHLAAAAYISDATWI 343
QY 229 GTALLPHQWQHKVHFVMSLHSHWVHAP-FRADHMLYECESFWAGSGRLVHGRVLRQ 286
DB 344 ETALRPHSKRGFIPSMALTLDHSISWMTDNFRVDDMLYENHSTIAGGSRSLIEGKLWTR 403
QY 287 DGLVAVTCAOEGVIRV---KPOVSESK 310
DB 404 DGRVLFSTTQEQALIRTAKNRPNASPAK 430

RESULT 15

G87701

acyl-CoA thiolesterase II [imported] - Caulobacter crescentus

C:Species: Caulobacter crescentus

C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001

C:Accession: G87701

R:Nietman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidele-
B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.;
N. J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser,
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A:Title: Complete Genome Sequence of Caulobacter crescentus.

A:Reference number: A87249; MUID:21173698; PMID:11259647

A:Accession: G87701

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-284 <STO>

A:Cross-references: GB:AE005673; NID:gl3425405; PIDN:AAK25611.1; GSPDB:GN00148

C:Genetics:

A:Gene: CC3649

C:Superfamily: acyl-CoA thiolesterase II

Query Match 30.1%; Score 501.5; DB 2; Length 284;

Best Local Similarity 41.5%; Pred. No. 6.5e-37;

Matches 118; Conservative 54; Mismatches 99; Indels 13; Gaps 9;

QY 23 VLNLEPLEDEDFRGRHYWVPAKRLFGGQIVGOALVAAAKSVSESVHSLHCYFVRAGD 82

DB 10 ILDELETVNLFRGVSPNDGPPRIFGGLVIAQALLAAVKTVPDRV-CHSLHAYFIRPGDV 68

QY 83 KPLVLYOVERTRTGSSFSVRSVKAQHGKPIFICQASFOQAQSPMOHQFSMTVPVPPPE 142

DB 69 TAPVLYEVRARDGGTFTTRVAAIQHGEQIFNLAASTQTPEDG-FEQSEMPASVDRES 127

QY 143 LDCETLIDQYLRDPLNOKRYPLALNRIRIAAOEVPTEIRKPNP-SPLSQORMPEKQMFV 201

DB 128 L---PTDAD-FLR--SLGDQIHKKVVAIAERPREVDIRWIDPQNPIAPVKSGTKQV-WM 180

QY 202 RARGYIGEGDMKHCVCVAAVYISDYAFGLTALLPHQ--WQHKVHFVMSLHSHWVHAPFRA 259

DB 181 RAKAPLGD-DVKQQAALAVASDNMFESALRPHGLIWTTPGICQAASLDHAMWFFHPPNF 239

QY 260 DHMMLYECESFWAGSGRLVHGRVLRQDGLVAVTCAOEGVIRV 303

DB 240 NDWTLFAQDPSASQGRGLVRGOMFSQDGKLLASVAQECLEMRV 283

Search completed: August 19, 2002, 13:19:52

Job time: 78 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 19, 2002, 13:19:34 ; Search time 13.48 Seconds
(without alignments)
893.307 Million cell updates/sec

Title: US-09-766-366-1

Perfect score: 1666

Sequence: 1 MGRAVATAALPPGDLRLV.....VTCAOEGVIRVPQVSESKL 311

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1621	97.3	319	1	PTL_HUMAN	O14734 homo sapien
2	1414.5	84.9	320	1	PTL_MOUSE	P58137 mus musculus
3	553	33.2	285	1	TESB_ECOLI	P23911 escherichia
4	460	27.6	286	1	TESB_HAEIN	P44498 haemophilus
5	273	16.4	349	1	PTL_YEAST	P41903 saccharomyc
6	99.5	6.0	417	1	IEG3_HSVSA	P13199 herpesvirus
7	93	5.6	354	1	Y101_NPVP	P24653 oryza psee
8	90	5.4	454	1	PYRC_METH	O27199 methanobact
9	90	5.4	2483	1	COR2_HUMAN	O00763 homo sapien
10	89.5	5.4	1483	1	CYP1_YEAST	P12351 saccharomyc
11	87	5.2	1654	1	PCFB_HUMAN	O94913 homo sapien
12	85.5	5.1	527	1	TXK_HUMAN	P42681 homo sapien
13	85.5	5.1	1070	1	PTK7_HUMAN	O13308 homo sapien
14	85	5.1	1068	1	HPR_HUMAN	O75146 homo sapien
15	84	5.0	648	1	CH44_YEAST	P43634 saccharomyc
16	84	5.0	697	1	TRP_SCHPO	O13831 schizosacch
17	83.5	5.0	529	1	UD16_RAT	P08430 rattus norv
18	83	5.0	816	1	QALF_NEUCR	P11638 neurospora
19	82.5	5.0	4466	1	DYRC_ANTCR	P39057 anthocidari
20	82	4.9	1585	1	P3K3_DICDI	P54675 dictyosteli
21	81.5	4.9	1177	1	MED_BACSU	P37474 bacillus su
22	81	4.9	344	1	CRM_CANFA	P79145 canis fami
23	80.5	4.8	591	1	DSM1_FSEAE	O9huw5 pseudomonas
24	79.5	4.8	1068	1	HPR_MOUSE	O9jky5 mus musculu
25	79.5	4.8	3511	1	MY15_MOUSE	O9qz24 mus musculu
26	79.5	4.8	5217	1	HTS1_COCCA	O01886 cochiobolu
27	79	4.7	251	1	GD80_WHEAT	P08079 triticum ae
28	79	4.7	316	1	HEMZ_DEIRA	O9rv38 deinococcus
29	79	4.7	364	1	IKBE_MOUSE	O54910 mus musculu
30	79	4.7	439	1	GFO_ZYMMO	O07982 zymomonas m
31	79	4.7	860	1	MUTS_PASMU	P57972 pasteurella
32	78.5	4.7	1350	1	NOS_DROME	O27571 drosophila
33	78	4.7	229	1	VG01_BPML5	O05218 mycobacteri

RESULT 1

ID	PTL_HUMAN	STANDARD	PRT	319 AA
AC	O14734; O15261;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	01-MAR-2002 (Rel. 41, Last annotation update)			
DE	Peroxisomal acyl-coenzyme A thioester hydrolase 1 (EC 3.1.2.2)			
DE	(Peroxisomal long-chain acyl-CoA thioesterase 1) (HIV-Nef associated acyl CoA thioesterase) (Thioesterase II) (hTE).			
GN	PTL1 OR HNAACTE.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=97445158; PubMed=9299485;			
RA	Watanabe H., Shiratori T., Shoji H., Miyatake S., Okazaki Y., Ikuta K., Sato T., Saito T.;			
RT	"A novel acyl-CoA thioesterase enhances its enzymatic activity by direct binding with HIV Nef."			
RL	Biochem. Biophys. Res. Commun. 238:234-239(1997).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Lymphoid;			
RX	MEDLINE=97298085; PubMed=9152323;			
RA	Liu L.X., Margottin F., LeGall S., Schwartz O., Selig L., Benarous R., Benichou S.;			
RT	"Binding of HIV-1 Nef to a novel thioesterase enzyme correlates with Nef-mediated CD4 down-regulation."			
RL	J. Biol. Chem. 272:13779-13785(1997).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Muscle;			
RX	MEDLINE=99194760; PubMed=10092594;			
RA	Jones J.M., Nau K., Geraghty M.T., Erdmann R., Gould S.J.;			
RT	"Identification of peroxisomal acyl-CoA thioesterases in yeast and humans."			
RL	J. Biol. Chem. 274:9216-9223(1999).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RA	Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R., Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baggeley C.L., Bailey J., Barlow K.F., Bates K.N., Bead L.M., Beare D.M., Beasley O.P., Bird C.P., Blakey S.E., Bridgman A.M., Brown A.J., Buck D., Burrill W., Butler A.P., Carder C., Carter N.P., Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M., Clegg S., Cobley V.E., Collier R.E., Connor R., Corby N.R., Coulson A., Coville G.J., Deadman R., Dhami P., Dunn M., Ellington A.G., Frankland J.A., Fraser A., French L., Garner P., Graham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E., Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J., Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D., Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S., LeVasialho M.H., Leversha M., Lloyd C., Lloyd D.M., Lovell J.D.,			

34	78	4.7	317	1	YVCK_BACSU	O06974 bacillus su
35	78	4.7	481	1	DOK1_HUMAN	Q99704 homo sapien
36	78	4.7	616	1	SPAS_HUMAN	Q9ubp0 homo sapien
37	77.5	4.7	304	1	GDB1_WHEAT	P04729 triticum ae
38	77.5	4.7	322	1	YQ16_LISIN	Q928c0 listeria in
39	77.5	4.7	778	1	TAST_HUMAN	O12815 homo sapien
40	77.5	4.7	812	1	AXN2_BRARE	P57095 brachydanio
41	77.5	4.7	817	1	HUNB_MUSDO	O01778 musca domes
42	77	4.6	291	1	Y499_MYCTU	Q11164 mycobacteri
43	77	4.6	609	1	FRI_ARATH	Q9fdw0 arabidopsis
44	77	4.6	875	1	FOCD_ECOLI	P46009 escherichia
45	77	4.6	993	1	EPB3_MOUSE	P54754 mus musculu

ALIGNMENTS

RA Marsh V.L., Martin S.L., McConnachie L.J., McLeay K., McMurray A.A.,
RA Milne S., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkneen R., Sims S.,
RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
RA Swann M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA Rogers J.;
RT "The DNA sequence and comparative analysis of human chromosome 20.":
RL Nature 414:865-871(2001).
CC -1- FUNCTION: MAY PLAY A ROLE IN FATTY ACID OXIDATION RATHER THAN
CC FORMATION OF FATTY ACIDS. MAY MEDIATE NEF-INDUCED DOWN-REGULATION
CC OF CDA.
CC -1- CATALYTIC ACTIVITY: Palmitoyl-CoA + H(2)O = CoA + palmitate.
CC -1- SUBUNIT: INTERACTS WITH HIV-1 NEF.
CC -1- SUBCELLULAR LOCATION: Peroxisomal.
CC -1- SIMILARITY: BELONGS TO THE C/M/P THIOESTER HYDROLASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF014404; AAB71665.1; -;
CC EMBL; X86032; CAAG0024.1; -;
CC EMBL; AF124284; AAD27616.1; -;
CC EMBL; AL008726; CAA15502.1; -;
CC HSP; P23911; 1680.
CC InterPro: IPR003703; Acyl-CoA:thio.
CC Pfam: PF02551; Acyl-CoA:thio; 1.
KW Hydrolyase; Serine esterase; Peroxisome.
FT ACT_SITE 78 78 BY SIMILARITY.
FT ACT_SITE 232 232 BY SIMILARITY.
FT SITE 317 319 MICROBODY TARGETING SIGNAL (POTENTIAL).
FT CONFLICT 291 293 LWR -> VMS (IN REF. 2).
FT CONFLICT 319 319 L -> R (IN REF. 2).
SQ SEQUENCE 319 AA; 35914 MW; 8345C6E5EABF3326 CRC64;

Query Match 97.3%; Score 1621; DB 1; Length 319;
Best Local Similarity 100.0%; Pred. No. 2.1e-138;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 PPGDLSRVLTVTNLEPLDEDLFRGRHYWVP-PAKRLFGGQIVGQALVAAAKSVSEDVH 70
Db 19 PPGDLSRVLTVTNLEPLDEDLFRGRHYWVP-PAKRLFGGQIVGQALVAAAKSVSEDVH 78
Qy 71 SLHCYFVRAGDPKPLVLYQVVERTGSSFSVRSVKAQVHGKPIFCOASFQQAQSPMQH 130
Db 79 SLHCYFVRAGDPKPLVLYQVVERTGSSFSVRSVKAQVHGKPIFCOASFQQAQSPMQH 138
Qy 131 QFSMPTVPPPEELLDCETLIDQYLRDPNLQKRYPLALNRIAAQEVPIETKVPNPSLSQL 190
Db 139 QFSMPTVPPPEELLDCETLIDQYLRDPNLQKRYPLALNRIAAQEVPIETKVPNPSLSQL 198
Qy 191 QRMPEKQMFWRVARGYIGEGDMKHCVCVAAVYISDYAFGLGTALLPHQWQHKVHFVMSLDHS 250
Db 199 QRMPEKQMFWRVARGYIGEGDMKHCVCVAAVYISDYAFGLGTALLPHQWQHKVHFVMSLDHS 258
Qy 251 MWFHAPPRADHMLYECESPWAGSGRGLVHGRWRDGLVAVTCAQEGVIRVKPQVSESK 310
Db 259 MWFHAPPRADHMLYECESPWAGSGRGLVHGRWRDGLVAVTCAQEGVIRVKPQVSESK 318

Qy 311 L 311
Db 319 L,319

RESULT 2
PTEL_MOUSE
ID PTEL_MOUSE STANDARD; PRT; 320 AA.
AC P58137; 2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Peroxisomal acyl-coenzyme A thioester hydrolase 1 (EC 3.1.2.2)
DE Peroxisomal long-chain acyl-coA thioesterase 1).
GN PTEL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: MAY PLAY A ROLE IN FATTY ACID OXIDATION RATHER THAN
CC FORMATION OF FATTY ACIDS (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: Palmitoyl-CoA + H(2)O = CoA + palmitate.
CC -1- SUBCELLULAR LOCATION: Peroxisomal (By similarity).
CC -1- SIMILARITY: BELONGS TO THE C/M/P THIOESTER HYDROLASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; BC005792; AAH05792.1; -;
CC InterPro: IPR003703; Acyl-CoA:thio.
CC Pfam: PF02551; Acyl-CoA:thio; 1.
KW Hydrolyase; Serine esterase; Peroxisome.
FT ACT_SITE 79 79 BY SIMILARITY.
FT ACT_SITE 233 233 BY SIMILARITY.
FT SITE 318 320 MICROBODY TARGETING SIGNAL (POTENTIAL).
SQ SEQUENCE 320 AA; 35827 MW; 94FGAFCFE2FA23 CRC64;

Query Match 84.9%; Score 1414.5; DB 1; Length 320;
Best Local Similarity 87.3%; Pred. No. 7.6e-120;
Matches 262; Conservative 22; Mismatches 15; Indels 1; Gaps 1;

Qy 13 GDLRSVLVTTNLEPLDEDLFRGRHYWVP-AKRLFGGQIVGQALVAAAKSVSEDVHVS 71
Db 21 GDLRSVLVTTNLEPLDEDLFRGRHYWVP-TSQLFGQIMGQALVAAAKSVSEDVHVS 80
Qy 72 LHCYFVRAGDPKPLVLYQVVERTGSSFSVRSVKAQVHGKPIFCOASFQQAQSPMQH 131
Db 81 LHCYFVRAGDPKPLVLYQVVERTGSSFSVRSVKAQVHGKPIFCOASFQQAQSPMQH 140
Qy 132 FSKPVPVPPPEELLDCETLIDQYLRDPNLQKRYPLALNRIAAQEVPIETKVPNPSLSQL 191
Db 141 FSKPVPVPPPEELLDCETLIDQYLRDPNLQKRYPLALNRIAAQEVPIETKVPNPSLSQL 200
Qy 192 RMEPKQMFWRVARGYIGEGDMKHCVCVAAVYISDYAFGLGTALLPHQWQHKVHFVMSLDHS 251
Db 201 ALEPKQMFWRVARGYIGEGDMKHCVCVAAVYISDYAFGLGTALLPHQWQHKVHFVMSLDHS 260
Qy 252 WFHAPPRADHMLYECESPWAGSGRGLVHGRWRDGLVAVTCAQEGVIRVKPQVSESK 311
Db 261 WFHAPPRADHMLYECESPWAGSGRGLVHGRWRDGLVAVTCAQEGVIRVKPQVSESK 320

RESULT 3
TESB_ECOLI
ID TESP_ECOLI STANDARD; PRT; 285 AA.
AC P23911;
DT 01-MAR-1992 (Rel. 21, Created)

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

EMBL; M63308; AAA24665.1; -;
 EMBL; AE000151; AAC73355.1; -;
 EMBL; U82664; AAB40208.1; -;
 EMBL; AE005224; AAG54802.1; -;
 EMBL; AP002551; BAB33929.1; -;
 PIR; JH0411; JH0411.
 PDB; 1C8U; 12-JUL-00.
 EcoGene; EG10995; tesb.
 InterPro; IPR003703; Acyl_CoA_thio.
 Pfam; PF02551; Acyl_CoA_thio.1.
 Hydrolase; 3D-structure; Complete proteome.
 INIT_MET 0
 FT ACT_SITE 57 57
 FT ACT_SITE 203 203
 SQ SEQUENCE 285 AA; 31835 MW; 6C761C7E9660D1EA CRC64;

Query Match 33.2%; Score 553; DB 1; Length 285;
 Best Local Similarity 42.1%; Pred. No. 1.8e-42;
 Matches 125; Conservative 53; Mismatches 91; Indels 28; Gaps 10;

QY 15 LRSVLVTVLNLEPDEDLFRGHRYWYPAKRLFGQIVGQALVAAAKSVSDVHVHSLRC 74
 Db 4 LKNLL--TLNLEKIEISLFRGQSEDGLGRQVGGVVGQALYAAKETVPEERLVHSFS 61
 QY 75 YFVRAGDPKLPVLVQVETRTTGSFSVRSKVAOHGKPIFCIQASFOQAQSPMQHFSM 134
 Db 62 YFLRPGSGSKPIIYDVETLRLDGNFSARRVAAIQNGKPIFYMTASF-QAPEAGFEHOKTM 120
 QY 135 PTVPPEPELDCETLIDQ---YLRLDPNLQKRYPLALNRIAAQVEPIEIKPVN-PSPLSQL 190
 Db 121 PSAPAPDG-LPSETQIAQSLAHLPPVLKDKF-----ICDRPLVRVPEFHNPLKG- 170
 QY 191 QRMPEPKQMWVRARGYIGEGDKMKHCVAAYISDYAFGLTALLPHOMQHKVHF- 244
 Db 171 HVAEPHQVWRIRANGSVYPD-DLRVHQVLLGYASDLNLPVALQ----HGTFLEPGIQI 225
 QY 245 VSLDHSWFFHAPRADHMLLYECESPWAGSGRLVHGRLRQDGLVATCAQEGVIR 301
 Db 226 ATIDHSWFFHPPNMLNELLVSVESASSARGFVRGEFTQDGLVASTVQEGVMR 282

RESULT 4
 TESB_HAEIN
 ID TESB_HAEIN STANDARD; PRT; 286 AA.
 AC P44436;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DE 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Acyl-CoA thioesterase II (EC 3.1.2.-) (TEII).
 OS TESB OR H10076.
 GN Haemophilus influenzae.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OC Haemophilus.
 OC NCBI_TaxID=727;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-RD / KW20 / ATCC 51907;
 RX MEDLINE=95350630; PubMed=7542800;
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 RA Kerkavane A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
 RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
 RA Fine L.B., Frichman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 RA Venter J.C.;
 RT "whole-genome random sequencing and assembly of Haemophilus

16-OCT-2001 (Rel. 40, Last annotation update)
 Dihydroorotase (EC 3.5.2.3) (DHOase).
 PYRC OR MTH1127.
 OS Methanobacterium thermoautotrophicum.
 CC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;
 CC Methanothermobacter.
 CC NCBI_TaxID=145262;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DELTA H;
 RX MEDLINE=98037514; PubMed=9371463;
 RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
 RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
 RA Harrison D., Hoang L., Keagle P., Lum W., Pothier B., Qiu D.,
 RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
 RA Jiwani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
 RA McDougall S., Shiner G., Goyal A., Pietrowski S., Church G.M.,
 RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
 RA "Complete genome sequence of Methanobacterium thermoautotrophicum
 deltaH: functional analysis and comparative genomics.";
 J. Bacteriol. 179:7135-7155(1997).
 RN [2]
 RP SEQUENCE OF 24-454 FROM N.A.
 RC STRAIN=DELTA H;
 RX MEDLINE=95014084; PubMed=7929010;
 RA Pihl T.D., Sharma S., Reeve J.N.;
 RT "Growth phase-dependent transcription of the genes that encode the
 two methyl coenzyme M reductase isoenzymes M and M5-
 methyltetrahydromethanopterin:coenzyme M methyltransferase in
 Methanobacterium thermoautotrophicum delta H.";
 J. Bacteriol. 176:6384-6391(1994).
 RL [1]
 CC -1- CATALYTIC ACTIVITY: (S)-dihydroorotate + H(2)O = N-carbamoyl-L-
 aspartate.
 CC -1- COFACTOR: THIS ENZYME TIGHTLY BINDS ONE ZINC ATOM PER CHAIN WHICH
 IS REQUIRED FOR THE CATALYTIC MECHANISM. IT ALSO BINDS WEAKLY TO
 TWO OTHER ZINCS WHICH ARE NOT ESSENTIAL FOR ACTIVITY
 (BY SIMILARITY).
 CC -1- PATHWAY: THIRD STEP IN PYRIMIDINE BIOSYNTHESIS.
 CC -1- SIMILARITY: BELONGS TO THE DHOASE FAMILY. SUBFAMILY 2.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AE000882; AAB85616.1; -;
 CC EMBL: U09990; AAA73440.1; -;
 CC MEROPS: M38.972; -;
 DR InterPro: IPR002195; Dihydroorotase.
 DR Pfam: PF00744; Dihydroorotase; 1.
 DR PROSITE: PS00482; DIHYDROOROTASE_1; 1.
 DR PROSITE: PS00483; DIHYDROOROTASE_2; 1.
 KW Pyrimidine biosynthesis; Hydrolase; Zinc; Complete proteome.
 FT METAL 78 78
 FT ZINC (POTENTIAL).
 FT METAL 80 80
 FT ZINC (POTENTIAL).
 FT CONFLICT 24 25 MF -> SL (IN REF. 2).
 FT CONFLICT 67 67 G -> A (IN REF. 2).
 FT CONFLICT 156 156 F -> N (IN REF. 2).
 FT CONFLICT 218 224 ARPPLAE -> CKTSRR (IN REF. 2).
 FT CONFLICT 265 266 CE -> SQ (IN REF. 2).
 FT CONFLICT 366 366 A -> P (IN REF. 2).
 FT CONFLICT 418 427 GPVMTIVRGR -> ARHDHCQG (IN REF. 2).
 SQ SEQUENCE 454 AA; 5053 MW; 40DE9F07B5D4BF83 CRC64;

Query Match 5.4%; Score 90; DB 1; Length 454;
 Best Local Similarity 22.7%; Pred. No. 1.3;
 Matches 48; Conservative 22; Mismatches 69; Indels 72; Gaps 8;
 QY 61 KVSVDVHVHSLHC-----YFVRAGDPKLPVLYQVETRTGTSFSFSRSVKA--VOHG 110

Db 181 KDLSDAPVPV-TVHCENRDVVMKSMKELKDRSDPSAYALARPPPLAEVSAEVLALSIIHHE 239
 QY 111 KPIFICQAFQQAQPSMQHQSPTVPPPEELDCETLIDQYLRDPLQKR----- 162
 Db 240 HPVHICHLSTVKA-----LQLVEPFREYVTCEVTPHLLLDGSAFRFRGTMVKTN 289
 QY 163 -----YPLALNRI-----AAQEVPIEIKPVNPSP----- 186
 Db 290 PPLRPPPSRIYPEFLDRINIGTDHAPHGTEKKRGWDAPPGIPNLEVLKLLTLVSK 349
 QY 187 -----LSQLQRM-----EPKQFWVRARGYIGEG 210
 Db 350 GRMSLSTIRMLAEAPARIEGLRSKGRIAG 380
 RESULT 9
 COA2_HUMAN
 ID COA2_HUMAN STANDARD; PRT; 2483 AA.
 AC Q00763; Q16852;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Acetyl-CoA carboxylase 2 (EC 6.4.1.2) (ACC-beta) [Includes: Biotin
 carboxylase (EC 6.3.4.14)].
 GN ACACB OR ACC2 OR ACCB.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.
 CC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=97256787; PubMed=9099716;
 RA Abu-Elheiga L., Almaraz-Ortega D.B., Baldini A., Wakil S.J.;
 RT "Human acetyl-CoA carboxylase 2. Molecular cloning, characterization,
 RT chromosomal mapping, and evidence for two isoforms.";
 J. Biol. Chem. 272:10669-10677(1997).
 RN [2]
 RP SEQUENCE OF 1349-2134 FROM N.A.
 RC TISSUE=Adipose tissue;
 RX MEDLINE=96265061; PubMed=8670171;
 RA Widmer J., Fassihi K.S., Schlichter S.C., Wheeler K.S., Crute B.E.,
 RA King N., Nuttle-Mcmenemy N., Noll W.W., Daniel S., Ha J., Kim K.-H.,
 RA Witters L.A.;
 RT "Identification of a second human acetyl-CoA carboxylase gene.";
 J. Biochem. J. 316:915-922(1996).
 CC -1- FUNCTION: ACC-BETA MAY BE INVOLVED IN THE PROVISION OF MALONYL-COA
 OR IN THE REGULATION OF FATTY ACID OXIDATION, RATHER THAN FATTY
 ACID BIOSYNTHESIS. THIS PROTEIN CARRIES THREE FUNCTIONS: BIOTIN
 CARBOXYL CARRIER PROTEIN, BIOTIN CARBOXYLASE, AND
 CARBOXYLTRANSFERASE.
 CC -1- CATALYTIC ACTIVITY: ATP + acetyl-CoA + HCO(3)(-) = ADP + phosphate
 + malonyl-CoA.
 CC -1- CATALYTIC ACTIVITY: ATP + biotin-carboxyl-carrier protein + CO(2)
 = ADP + phosphate + carboxybiotin-carboxyl-carrier protein.
 CC -1- COFACTOR: BIOTIN.
 CC -1- PATHWAY: FIRST STEP (RATE LIMITING) IN LONG-CHAIN FATTY ACID
 SYNTHESIS.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC; MAY ASSOCIATE WITH MEMBRANES.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
 SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN THE HEART, SKELETAL
 MUSCLES AND LIVER.
 CC -1- SIMILARITY: PARTIAL TO CARBAMOYL PHOSPHATE SYNTHETASES.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----

RESULT	ID	CYPL_YEAST	STANDARD;	PRT;	1483 AA.
CYPL_YEAST	ID	CYPL_YEAST	STANDARD;	PRT;	1483 AA.
AC	P12351;				
DDT	01-OCT-1989	(rel. 12, Created)			
DDT	01-OCT-1989	(rel. 12, Last sequence update)			
DDT	01-MAR-2002	(rel. 41, Last annotation update)			

CYP1 activatory protein.
 CYP1 OR HAP1 OR YLR256W.
 Saccharomyces cerevisiae (Baker's yeast).
 Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 NCBI_Taxid=4932;
 [1]
 SEQUENCE FROM N.A.
 MEDLINE=89125585; PubMed=2851659;
 CREUSOT F., Verdier J., Galsne M., Slonimski P.P.;
 "CYP1 (HAP1) regulator of oxygen-dependent gene expression in yeast.
 I. Overall organization of the protein sequence displays several
 novel structural domains.";
 J. Mol. Biol. 204:263-276(1988).
 [2]
 SEQUENCE FROM N.A.
 MEDLINE=89106221; PubMed=2643482;
 Preifer K., Kim K.-S., Kogan S., Guarente L.;
 "Functional dissection and sequence of yeast HAP1 activator.";
 Cell 56:291-301(1989).
 [3]
 MUTANT CYP1-18.
 MEDLINE=89125586; PubMed=2851659;
 Verdier J., Galsne M., Guard B., Defranoux N., Slonimski P.P.;
 "CYP1 (HAP1) regulator of oxygen-dependent gene expression in yeast.
 II. Missense mutation suggests alternative Zn fingers as
 discriminating agents of gene control.";
 J. Mol. Biol. 204:277-282(1988).
 [4]
 DNA-BINDING.
 MEDLINE=97042358; PubMed=8887558;
 Zhang L., Guarente L.;
 "The C6 zinc cluster dictates asymmetric binding by HAP1.";
 EMBO J. 15:4676-4681(1996).
 [5]
 STRUCTURE BY NMR OF 60-100
 MEDLINE=96275662; PubMed=8683583;
 Zimmerman J., Vuidepot A.-L., Bontems F., Lallemand J.-Y., Gervais M.,
 Flechter B., Guard B.;
 "1H, 15N resonance assignment and three-dimensional structure of CYP1
 (HAP1) DNA-binding domain.";
 J. Mol. Biol. 259:752-804(1996).
 [6]
 X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 55-135 OF MUTANT HAP1-l8
 COMPLEXED WITH DNA.
 MEDLINE=99101378; PubMed=9886287;
 King B.A., Zhang L., Guarente L., Marmorstein R.;
 "Structure of HAP1-l8-DNA implicates direct allosteric effect of
 protein-DNA interactions on transcriptional activation.";
 Nat. Struct. Biol. 6:22-27(1999).
 [7]
 X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 56-135 IN COMPLEX WITH DNA.
 MEDLINE=99101385; PubMed=9886294;
 King D.A., Zhang L., Guarente L., Marmorstein R.;
 "Structure of a HAP1-DNA complex reveals dramatically asymmetric DNA
 binding by a homodimeric protein.";
 Nat. Struct. Biol. 6:64-71(1999).
 [8]
 X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 55-135 OF MUTANT HAP1-PC7
 COMPLEXED WITH DNA.
 MEDLINE=20482201; PubMed=11024163;
 Lukens A.K., King D.A., Marmorstein R.;
 "Structure of HAP1-PC7 bound to DNA: Implications for DNA recognition
 and allosteric effects of DNA-binding on transcriptional activation.";
 Nucleic Acids Res. 28:3853-3863(2000).
 -1- FUNCTION REGULATION OF OXYGEN DEPENDENT GENE EXPRESSION. IT
 MODULATES THE EXPRESSION OF ISO-1 AND ISO-2 CYTOCHROME C. BINDS
 TO THE SEQUENCE 5'-CGGNNWTCGG-3'.
 -1- SUBUNIT. BINDS DNA AS A HOMODIMER.
 -1- SUBCELLULAR LOCATION: Nuclear.
 -1- MISCELLANEOUS: CYP1-18 MUTANT ACTIVATES THE EXPRESSION OF CYP3
 (ISO-2) WHILE REDUCING THAT OF CYP1 (ISO-1).
 -1- MISCELLANEOUS: HEME IS AN EFFECTOR MOLECULE FOR CYP1/HAP1. THE


```
Query Match 5.2%; Score 87; DB 1; Length 1654;
Best Local Similarity 21.1%; Pred. No. 11;
Matches 52; Conservative 33; Mismatches 70; Indels 92; Gaps 10;

QY 1 MGRAVATAALPPGDLRSVLTTLNLEPLED-----LFRGRHYW---VPAKRLFGGOIVG 53
DB 183 VGREYLTAF-----TKMLVATFICVKEKVDENTRKSFLKURSTWDEIFPLKKLY----- 231
QY 54 QALVAAAKSVSDVHSHLCYFVRAGDKPLVLYQVETRTGSSFSVRSKAVQHGKPI 113
DB 232 -----ALDVRVNSL-----DPAMPKPLPPNVTSSI-----HVNPK 263
QY 114 FICOASFOQAQSPMQHFSM---PVPVPEELLDCETLIDQYL---RDPNL-----QKRYP 164
DB 264 FLNKSPEEPSTGVVSSPSISTPPIVDIQKNLTQQLIRQQLAKQKOLLELQOKKLE 323
QY 165 LALNRTAAQ-----EVPTEIKPVNPSPL 187
DB 324 LELEQAQALAVLSVQOETSNLPGSAPSKLHVSIQIPPMAYKAPHQVPVQSEKSRGGS 383
QY 188 SQJORME 194
DB 384 LQIOTDLK 390

RESULT 12
TXK_HUMAN STANDARD; PRT; 527 AA.
ID TXK_HUMAN Q14220;
AC P42881; Q14220;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Tyrosine-protein kinase TXK (EC 2.7.1.112).
GN TXK.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Blood;
RX MEDLINE=95038742; PubMed=7951233;
RA Haire R.N., Ohta Y., Lewis J.E., Fu S.M., Kroisel P.M., Litman G.W.;
RT "TXK, a novel human tyrosine kinase expressed in T cells shares
RT sequence identity with Tec family kinases and maps to 4p12.";
RL Hum. Mol. Genet. 3:897-901(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-Blood;
RX MEDLINE=96197775; PubMed=8632917;
RA Ohta Y., Haire R.N., Amemiya C.T., Litman R.T., Trager T., Riess O.,
RA Litman G.W.;
RT "Human Txk: genomic organization, structure and contiguous physical
RT linkage with the Tec gene.";
RL Oncogene 12:937-942(1996).
CC -/- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -/- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -/- TISSUE SPECIFICITY: EXPRESSED IN T CELLS AND SOME MYELOID CELL
CC LINES.
CC -/- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC -/- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -/- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. BTK
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
```

or send an email to license@isb-sib.ch).

```
CC EMBL; L27071; AAA74557.1; -
DR EMBL; U34379; AAB60412.1; -
DR EMBL; U34367; AAB60412.1; JOINED.
DR EMBL; U34368; AAB60412.1; JOINED.
DR EMBL; U34359; AAB60412.1; JOINED.
DR EMBL; U34370; AAB60412.1; JOINED.
DR EMBL; U34371; AAB60412.1; JOINED.
DR EMBL; U34372; AAB60412.1; JOINED.
DR EMBL; U34373; AAB60412.1; JOINED.
DR EMBL; U34374; AAB60412.1; JOINED.
DR EMBL; U34375; AAB60412.1; JOINED.
DR EMBL; U34376; AAB60412.1; JOINED.
DR EMBL; U34377; AAB60412.1; JOINED.
DR EMBL; U34378; AAB60412.1; JOINED.
DR HSSP; Q06187; IAWX.
DR MIM; 600058; -
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR000980; SH2.
DR InterPro; IPR001452; SH3.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00017; SH2; 1.
DR Pfam; PF00018; SH3; 1.
DR PRINTS; PR00401; SH2DOMAIN.
DR PRINTS; PR00109; TYRKINASE.
DR SMART; SM00252; SH2; 1.
DR SMART; SM00326; SH3; 1.
DR SMART; SM00219; TYRK; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS50001; SH2; 1.
DR PROSITE; PS50002; SH3; 1.
KW Transferase; Tyrosine-protein kinase; ATP-binding; SH2 domain;
KW SH3 domain; Phosphorylation. POLY-CYS.
FT DOMAIN 14 19 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN 68 73 SH2.
FT DOMAIN 82 142 SH3.
FT DOMAIN 150 246 SH2.
FT DOMAIN 271 527 PROTEIN KINASE.
FT NP_BIND 277 285 ATP (BY SIMILARITY).
FT BINDING 299 299 ATP (BY SIMILARITY).
FT ACT_SITE 390 390 BY SIMILARITY.
FT MOD_RES 420 420 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT CONFLICT 45 45 H -> R (IN REF. 2).
SQ SEQUENCE 527 AA; 61239 MW; BC5D0705E4553003 CRC64;

Query Match 5.1%; Score 85.5; DB 1; Length 527;
Best Local Similarity 19.9%; Pred. No. 3.8; Mismatches 43; Indels 105; Gaps 16;
Matches 62; Conservative 43; Mismatches 102; Indels 105; Gaps 16;

QY 89 OVERTRTGSSFSVRS---VKAVOHGKPIFCQASFO-----QAQSPMQHOFMPTVP 139
DB 22 OKRQMTQISLSDDELPKTYTOHRRP-WLSQLSNKKQNTGRVQPSKRRK---PLPLPP 77
QY 140 ---PELLDCETLIDQYLRDP-----NLQKRYP-----LALNRI 170
DB 78 SEVAEEKIQVKALDYFLPREPCNLARRAEYLLILEKYNPHWKKADRLGNGLIPSNYV 137
QY 171 AAQEVPIEI-----KPVNPSPLSLQRMPEKQ-MFWVRARGVIGSGDKMKHCCVAAYS 223
DB 138 TENKINTLEIYEWYHRNITRNQAEHLRLQESKEGAFIVDRSLHSGYTISV----- 188
QY 224 DYAFICT-----ALLPH-----QWQ-HKVFHWVSLDHSKWFH----- 254
DB 189 ---FMGARSTAAIKYQIKKNDSCQWYVAERHAFQSIPELTIWYHQHNAAGLMTLRY 245
QY 255 -----APFRADHNMWLECESPW-----AGGSRGLVHGRWLRQDGLAVTCAQEGV 299
DB 246 VGLMGSCLPATAGFSYKEKWEIDPSELAFKEIGSGQFGVVHLGEWSHIQVAIKAINES 305
```

```

Qy 300 IRVKPQVSESKL 311
: : : : :
Db 306 MSEEDFIEAKV 317

RESULT 13
PTK7_HUMAN STANDARD; PRT; 1070 AA.
ID PTK7_HUMAN Q13417;
AC Q13308; Q13417;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Tyrosine-protein kinase-like 7 precursor (Colon carcinoma kinase-4)
GN (CKK-4).
GE PTK7 OR CKK4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
TISSUE=Colon carcinoma, and Placenta;
RX MEDLINE=96074849; PubMed=7478540;
RA Mossie K., Jallat B., Alves F., Sures I., Plozman G.D., Ullrich A.;
RT "Colon carcinoma kinase-4 defines a new subclass of the receptor
RT tyrosine kinase family."
RL Oncogene 11:2179-2184(1995).
[2]
SEQUENCE FROM N.A.
TISSUE=Fibroblast;
RX MEDLINE=97037064; PubMed=8882711;
RA Park S.-K., Lee H.-S., Lee S.-T.;
RT "Characterization of the human full-length PTK7 cDNA encoding a
RT receptor protein tyrosine kinase-like molecule closely related to
RT chick Klg."
RL J. Biochem. 119:235-239(1996).
CC -1- FUNCTION: MAY FUNCTION AS A CELL ADHESION MOLECULE. LACKS PROBABLY
CC THE CATALYTIC ACTIVITY OF TYROSINE KINASE. MAY BE CONNECTED TO THE
CC PATHOPHYSIOLOGY OF COLON CARCINOMAS AND/OR MAY REPRESENT A TUMOR
CC PROGRESSION MARKER.
CC -1- TISSUE SPECIFICITY: Type I membrane protein.
CC KIDNEY, PLACENTA AND MELANOCYTES. WEAKLY EXPRESSED IN THYROID
CC GLAND, OVARY, BRAIN, HEART AND SKELETAL MUSCLE. ALSO EXPRESSED IN
CC ERYTHROLEUKEMIA CELLS. BUT NOT EXPRESSED IN COLON.
CC -1- SIMILARITY: CONTAINS 7 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC -1- SIMILARITY: BELONGS TO THE INSULIN RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES. TRK-TYPE SUBFAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL: U33635; AAA87565.1; -
DR EMBL: U40271; AAC50484.1; -
DR HSP; P12931; IFMK.
DR MIM; 601890; -
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR003600; Ig-like.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00047; Ig; 7.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR SMART; SM00410; IG_Like; 2.
DR SMART; SM00408; IGG2; 5.
DR SMART; SM00219; TyrcK; 1.

```

PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 KW Receptor; Transmembrane; Signal; Glycoprotein; Cell adhesion;
 KW Immunoglobulin domain; Repeat.
 FT SIGNAL 1 30 POTENTIAL.
 FT CHAIN 31 1070 TYROSINE-PROTEIN KINASE-LIKE 7.
 FT DOMAIN 31 704 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 705 725 POTENTIAL.
 FT DOMAIN 726 1070 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 143 208 IG-LIKE C2-TYPE DOMAIN 1.
 FT DOMAIN 239 308 IG-LIKE C2-TYPE DOMAIN 2.
 FT DOMAIN 336 398 IG-LIKE C2-TYPE DOMAIN 3.
 FT DOMAIN 436 488 IG-LIKE C2-TYPE DOMAIN 4.
 FT DOMAIN 537 577 IG-LIKE C2-TYPE DOMAIN 5.
 FT DOMAIN 671 1066 IG-LIKE C2-TYPE DOMAIN 6.
 FT DOMAIN 796 1066 IG-LIKE C2-TYPE DOMAIN 7.
 FT DISULFID 53 101 PROTEIN KINASE; INACTIVE.
 FT DISULFID 150 200 BY SIMILARITY.
 FT DISULFID 246 301 BY SIMILARITY.
 FT DISULFID 343 391 BY SIMILARITY.
 FT DISULFID 433 481 BY SIMILARITY.
 FT DISULFID 524 570 BY SIMILARITY.
 FT DISULFID 613 664 BY SIMILARITY.
 FT CARBOHYD 116 116 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 175 175 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 184 184 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 214 214 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 268 268 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 283 283 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 405 405 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 463 463 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 567 567 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 646 646 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 92 92 P -> R (IN REF. 2).
 FT CONFLICT 147 147 K -> T (IN REF. 2).
 FT CONFLICT 207 207 S -> G (IN REF. 2).
 FT CONFLICT 495 496 VL -> RV (IN REF. 2).
 FT CONFLICT 515 515 G -> E (IN REF. 2).
 FT CONFLICT 834 834 S -> T (IN REF. 2).
 FT CONFLICT 881 881 E -> G (IN REF. 2).
 FT CONFLICT 969 969 A -> P (IN REF. 2).
 FT CONFLICT 992 992 S -> F (IN REF. 2).
 SQ SEQUENCE 1070 AA; 118260 MW; 47CDF25B8E3698A5 CRC64;

Query Match 5.1%; Score 85.5; DB 1; Length 1070;
 Best Local Similarity 21.6%; Pred. No. 9.1;
 Matches 72; Conservative 37; Mismatches 105; Indels 119; Gaps 18;

Qy 1 MGRAVATAALPPG-DLRSVLVTTVLN-----LEPLDELPRGRHYWPAKRLFGQ 50
 Db 1 MGAARSPARPRRLPLLSVLLPLLGQTQTATVFIKQPSQDALQGR----- 47
 Qy 51 IVGQALVAAKSVSDVHVHSLHCYFVRAGDPKLPVLYQVTRTSGSFSVRSKAVOH- 109
 Db 48 ---RALLRCEVEAPGPVHV-----YMLLDG---APVQDTERFAAGSSLSFAAVDPLQDS 96
 Qy 110 -----GKPIFICOASFQ---OAQSPMOHQFSMTVPVPPPEEL-LDCETLID- 151
 Db 97 GTFOCVARDVDVTGEARSANASFNKIEWIAGPVVVKHPASEAEIQPOTVKLRCH--IDG 154
 Qy 152 -----QYLRDPNLQKRYPL-----ALNRIAAQVEPIBIKVPVNPSPLSQLORMPEKQMFV 201
 Db 155 HPRPTYQWFRDGT-----PLSDGQSNHTVSSKERNLT/LRPAGP----- 192
 Qy 202 RARGYIGEDGMKHCVCV-AAY-----ISDYVFLGTALLPHQWQHKVHEVSLD 248
 Db 193 -----EHSGLYSCCAHSAFSAOCSNFTLSIADESFARVVLAPOD-----VVVARY 239
 Qy 249 HSMWFHAFPRAD-----HWMLYECESFWAGGSR 276
 Db 240 EEMAFHCQFSAQPPPSLQW-LFEDETPTNRSR 271

KW Actin-binding; Coiled coil.
FT DOMAIN 347 599 COILED COIL (POTENTIAL).
FT DOMAIN 800 1068 TALIN-LIKE.
SQ SEQUENCE 1068 AA; 119388 MW; 3CB7CF1191BFF8F CRC64;

Query Match 5.1%; Score 85; DB 1; Length 1068;
Best Local Similarity 19.0%; Pred. NO. 10;
Matches 72; Conservative 47; Mismatches 123; Indels 136; Gaps 14;

QY 14 DLRSVLTVTNLNLEPLD-----EDLFRGRHYWVPAKRLFGGQIVGOALVAAAKSVSEDV 67
Db 168 DVNNIFQLTVMEDYMOCELKLSVFRQLNTAIVSQMSSGGCRPLAPLIQVQDCSHLY 227
QY 68 H-----VHSLHCYFVRAGDPKLPVLY-----QVERTR 94
Db 228 HYTVKLLFKLHSCLPADTLQGHDRFHEQFHSLSRNFERRASD-----MLYFKRLIQLPRLP 283
QY 95 TGSSEFSVRSKAVQHGKIFIC-QASFOQAQSPMQHQFSMPVTPPEELDDCETLIDQY 153
Db 284 EGPPNFLRASALAEHIKPVVPEEAPDEEP---ENLIEISTGPPAGEPVVADLFDQT 340
QY 154 LRDPN-----LQ-----KRYPLALNRIRAAQEVPIETKPV----- 182
Db 341 FGPPNGSVKDDDDLDLQIESLKREVMRLRSELEKIKLEAQRVIAQLKSQVNALEGELEQK 400
QY 183 -----NPSPLSQLORMEPMKQMFVWRANGYIGEGDMKMHCCVAAVYISDYAFGLTALLP 234
Db 401 QKQKALVDNEOLRHLEAQLRAAQLEGERSQGLREAEERKASATEARY----- 447
QY 235 HQWQHKVHFVSLDHSNMFHAPFRADHMLYECESPWAGGSRGLVHGRLRQDG---VL 290
Db 448 NKLREK-----HSELVH-----VHAELLKKNADTAKQL 475
QY 291 AVT-CAQEGVIRVKPQVS 307
Db 476 TVTQSQEEVARVKEOLA 493

RESULT 15
CH4A_YEAST STANDARD; PRT; 648 AA.
ID CH4A_YEAST
AC P43634;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE CH4A activatory protein.
GN CH4A OR YLR098C OR L8004.8.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RA Schjerling P.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z.,
RA Javello A., Fulton L., Gattung S., Greco T., Kirsten J., Kucaba T.,
RA Hallsworth K., Hawkins J., Hillier L., Jier M., Johnson D.,
RA Johnston L., Langston Y., Latreille P., Le T., Mardis E., Menezes S.,
RA Miller N., Nhan M., Pauley A., Peluso D., Rifken L., Riles L.,
RA Taich A., Trevasakis E., Vignati D., Wilcox L., Wohlman P., Vaudin M.,
RA Wilson R., Waterston R.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
[3]
RN SEQUENCE FROM N.A.
RP Benes V., Rechmann S., Nentwich U., Schwager C., Ansoerge W., Voss H.;
RA Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
RL
CC -!- FUNCTION: ACTIVATES THE CH4A GENE FOR L-SERINE DEHYDRATASE. BINDS

RESULT 14
HIPR_HUMAN STANDARD; PRT; 1068 AA.
ID HIPR_HUMAN
AC Q75146; Q9UED9;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Huntingtin interacting protein 1 related (Hipl-related) (Hip 12).
GN HIP1L OR HIP12 OR KIAA0655.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20515263; PubMed=11063258;
RA Chopra V.S., Metzler M., Rasper D.M., Engqvist-Goldstein A.E.Y.,
RA Singaraja R., Gan L., Fichter K.M., McCutcheon K., Drubin D.,
RA Nicholson D.W., Hayden M.R.;
RL HIP12 is a non-proapoptotic member of a gene family including HIP1,
RL an interacting protein with huntingtin.;
RL Mamm. Genome 11:1006-1015(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=98403880; PubMed=9734811;
RA Ishikawa K.-I., Nagase T., Suyama M., Miyajima N., Tanaka A.,
RA Kotani H., Nomura N., Ohara O.;
RL "Prediction of the coding sequences of unidentified human genes. X.
RL The complete sequences of 100 new cDNA clones from brain which can
RL code for large proteins in vitro.";
RL DNA Res. 5:169-176(1998).
RN [3]
RP SEQUENCE OF 179-1068 FROM N.A.
RC TISSUE=Neuroblastoma;
RX MEDLINE=99069783; PubMed=9852681;
RA Seki N., Muramatsu M., Sugano S., Suzuki Y., Nakagawara A., Ohhira M.,
RA Hayashi A., Hori T., Saito T.;
RL "Cloning, expression analysis, and chromosomal localization of HIP1L,
RL an isolog of huntingtin interacting protein (HIP1).";
RL J. Hum. Genet. 43:268-271(1998).
CC -!- FUNCTION: COMPONENT OF CLATHRIN-COATED PITS AND VESICLES, THAT MAY
CC LINK THE ENDOCYTIC MACHINERY TO THE ACTIN CYTOSKELETON.
CC -!- SUBUNIT: INTERACTS WITH ACTIN AND HUNTINGTIN INTERACTING PROTEIN 1
CC (HIP1). DOES NOT INTERACT WITH HUNTINGTIN (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC. MEMBRANE-ASSOCIATED PROTEIN,
CC MAINLY LOCALIZED AT THE ENDOCYTIC COMPARTMENTS AND IN THE
CC PERINUCLEAR REGION (BY SIMILARITY).
CC -!- TISSUE SPECIFICITY: BRAIN, HEART, KIDNEY, PANCREAS, AND LIVER, BUT
CC NOT IN LUNG OR PLACENTA.
CC -!- DOMAIN: THE TALIN-LIKE DOMAIN BINDS F-ACTIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE SLA2 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB014555; BAA31630.1; AUT_INIT.
CC EMBL; AB013384; BAA33713.1; -
CC
CC MIM: 605613; -
CC InterPro: IPR001026; ENTH.
CC InterPro: IPR002558; ILWEQ.
CC Pfam: PF01417; ENTH; 1.
CC Pfam: PF01608; I_ILWEQ; 1.
CC ProDom: PD011820; ILWEQ; 1.
CC SMART: SM00273; ENTH; 1.
CC SMART: SM00307; ILWEQ; 1.

```
CC      TO THE DNA SEQUENCE 5'-GVGGARAYRTRATTCRC-3'.
CC      -!- SUBCELLULAR LOCATION: Nuclear.
CC      -!- SIMILARITY: CONTAINS 1 ZN(2)-CYS(6), FUNGAL-TYPE BINUCLEAR
CC      CLUSTER DOMAIN.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; Z49975; CAA90276.1; -
DR      EMBL; U53876; AAB67542.1; -
DR      EMBL; Z73270; CAA97662.1; -
DR      TRANSFAC; T02848; -
DR      SGD; S0004088; CHA4.
DR      InterPro; IPR001138; ZN2_Cy6_fungal.
DR      Pfam; PF00172; Zn_c1us; 1.
DR      SMART; SM00066; GAL4; 1.
DR      PROSITE; PS00463; ZN2_Cy6_FUNGAL_1; 1.
DR      PROSITE; PS50048; ZN2_Cy6_FUNGAL_2; 1.
KW      Transcription regulation; Activator; DNA-binding; Nuclear protein;
KW      Zinc; Metal-binding.
FT      DOMAIN 32 35 POLY-ASN.
FT      DNA_BIND 44 70 ZN(2)-CYS(6), FUNGAL-TYPE.
FT      DOMAIN 48 52 POLY-ARG.
FT      DOMAIN 105 109 POLY-SER.
FT      DOMAIN 195 198 POLY-GLN.
FT      DOMAIN 302 306 POLY-SER.
SQ      SEQUENCE 648 AA; 74392 MW; CF3381B4C0356F82 CRC64;

Query Match          5.0%; Score 84; DB 1; Length 648;
Best Local Similarity 21.9%; Pred. No. 6.7;
Matches 48; Conservative 32; Mismatches 49; Indels 90; Gaps 14;

QY 131 QFSMTVPPELLDCELTLDYLRDPN--LOKRYPL----ALNRIAAQEVPIET---KP 181
Db 394 RLSNSTVPETDELPEIETGIEEYIYDPKVLSTANPLKKLIIVLSRIT-----EIFASKI 447
QY 182 VNPSPLSQLQRM-----PKMFVWRARGYIGEGDMKMHCCVAAAYISDY 225
Db 448 FSPNE-TLLORSEYLAKFNLEVNWRDLPPELQWTK-----RSLMEM-----TDF 492
QY 226 AFLGTALLPHQVHVKVFWSLDHSWTFHAPRADHMLYCECEPWAGSGRLVHG---- 281
Db 493 ----NPTIAYVWFHYIVLIS-----YNKPF-----IYETKQ-----SRELVEGVDE 531
QY 282 -----RLWRQ-----DGLAVTCAQEGVIR 301
Db 532 LYLLKVMKNKFKTEKATYIMYISAILAIQCMKSNLIK 570
```

Search completed: August 19, 2002, 13:23:28
Job time: 234 sec

Db 1 MSRPETAT--PMD-----ALLAILLEKLEENLFRGLSPQVGNQVRFGGQVGOALVAAQ 54
Qy 61 KVSDEVHVHSLHCYFVRAGDKPLVLYOVERTTRTSSFSVRSVKAQVHGKPIFCOASF 120
Db 55 RTVDGGRVYVHSLHAFMRGDSVPYIYEVDRIDGSSFATRRVVAIQHKAIFAMASAF 114
Qy 121 QAAQSPMOHQFSMPTVPPPEELDCETLIDQYLDPNLOKRYPLALNRIAAAQEVPIEK 180
Db 115 QYDEDEG-FEHQDMPDVPMPETLPGQELKEKFL-----VHAPEAIRRYWERPRPIEIR 167
Qy 181 PVNPSPLSOLORMEPKQFVWRVARGVIGEGDKMKHCCVAAIYSDYAFGLTALLPH---QW 237
Db 168 PVLSHEYSRAKASQDQVWVAVGTPD-ERHQAVALYLSDMTLLDTSFYAHGTSVY 226
Qy 238 QHKVHFVSLDSHMFHAPFRADHWNLYECESPWAGSGRLVHGRWQDGVLAVTCAQE 297
Db 227 DRSLOQ-VASLDHAMFHRPSKMDWLLTYQDSPSAHGARGMTGSLFDRSGVLIASVAQE 285
Qy 298 GVIRVK 303
286 GLIRKK 291

RESULT 2

Q9HX74 PRELIMINARY; PRT; 289 AA.
AC Q9HX74;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DE ACYL-COA THIOESTERASE II.
GN TESB OR PA3942.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae.
OC Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
opportunistic pathogen.";
Nature 406:959-964(2000).
EMBL: AE004812; AAC07329.1; -.
DR HSSP; P23911; IC8U.
DR InterPro; IPR003703; Acyl_CoA_thio.
DR Pfam; PF02551; Acyl_CoA_thio; 1.
KW Complete proteome.
SQ SEQUENCE 289 AA; 32934 MW; 6592E1FFCE32DFE0 CRC64;

Query Match 34.1%; Score 568; DB 16; Length 289;
Best Local Similarity 42.1%; Pred. NO. 9.7e-49;
Matches 123; Conservative 48; Mismatches 95; Indels 26; Gaps 7;

Qy 20 VTTVLNLEPDEDLFRGRHYWVPAKRLFGGQIVGOALVAAKSVSEDEVHSLHCYFVRA 79
Db 8 LVALLSLEPIEENLFRGMSQDGLFQQLYGGVGLGALSAAQSTVEDRIVHSHMGYFLRP 67
Qy 80 GDPKPLVLYOVERTTRTSSFSVRSVKAQVHGKPIFCOASFQAAQSPMOHQFSMPTVPP 139
Db 68 GDASMPVYQVDRDGSFSTRRTAIOKQPIFLSSSFQYDEEG-FHHQIEMPEVVG 126
Qy 140 PEELDCETL-----IDQYLDPNLOKRYPLALNRIAAAQEVPIEKPVNPSPLSOLQR 192
Db 127 PDNLPSLELTRQAEIRIPERIRKLL---YP-----KPIEIRPVTOHQDPDPOP 173

Qy 193 MEPKQFVWRVARGVIGEGDKMKHCCVAAIYSDYAFGLTALLPHQ---QHKVHFVMSLDH 249
Db 174 DEPVKYLWFRADGKLPDVP-ALHKYLLAYASDFGLTTSMPHAYSVNQ-KTMQVASLDH 231
Qy 250 SMWFHAPFRADHWNLYECESPWAGSGRLVHGRWQDGVLAVTCAQEGVIR 301
Db 232 ALMFHRDLRADEWLLYIDAIDSPWAGNARGFVRGSIFFNQSGLVASSSQEGLIR 283
RESULT 3
Q19781 PRELIMINARY; PRT; 343 AA.
ID Q19781;
AC Q19781;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DE ACYL-COA THIOESTERASE II.
GN F25E2.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX Leimbac D., Minx M.;
RT "The sequence of C. elegans cosmid F25E2.";
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX Waterston R.;
RT "Direct Submission.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; U50197; AAA91256.1; -.
DR HSSP; P23911; IC8U.
DR InterPro; IPR003703; Acyl_CoA_thio.
DR Pfam; PF02551; Acyl_CoA_thio; 1.
KW Hypothetical protein.
SQ SEQUENCE 343 AA; 38845 MW; 239065057C69E8A0 CRC64;

Query Match 34.0%; Score 567; DB 5; Length 343;
Best Local Similarity 38.4%; Pred. NO. 1.5e-48;
Matches 126; Conservative 52; Mismatches 104; Indels 46; Gaps 9;

Qy 14 DLRSVLVTVLNLEPDEDLF-----RGRHYWVPAKRLFGGQIVGOALVAAKSVSEDEV 67
Db 27 DIKAGLIDTFLNLQRIDTTLNLYTARHLKGRHSY---NAVYGGVVGSLAAAAATVDCF 83
Qy 68 HVHSLHCYFVRAGDKPLVLYOVERTTRTSSFSVRSVKAQVHGKPIFCOASFQ----- 121
Db 84 IPHSLHSYFIKTSVYDKPKPLYMIDRIDGSRFCTRVVKAQVDCGEAIFSCQISFHHVCGN 143
Qy 122 -----QAQSPMOHQFSMPTVPPPEELDCETLIDQYLDPNLQ----- 160
Db 144 GTENKKRDTFOKEPDAIKHSSKMPVTPPEQLLPAREAALEVLRTKEVDEVTAQVIOHFL 203
Qy 161 KRYPLALNRIAAAQEVPIEKPVNPSPLSOLORMEPKQFVWRVARGVIGEGDKMKHCCVAA 220
Db 204 KEIPDAFERV-----FQVRPVNPAKYLKEDTEPMSMIWIRARENLGD-DHRHQCVAA 256
Qy 221 YISDYAFGLTALLPHQVHKVHFV-VSLDSHSMFHA-PFRADHWNLYECESPWAGSGRL 278
Db 257 YLTDLMLTAVRPHIRNGFIPMSFSLOHCIMWHENEFIDDMMLYETISSKAGGSRAF 316

[illegible]

RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furian L.R.,
RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Hoheisel J.D., Junqueira M.L.Z., Kemper E.L., Kitajima J.P.,
RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nhani A.J., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Peixoto B.R., Pereira G.A.G., Pereira H.M.Jr., Pesquero J.B.,
RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tshako M.H.,
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Weidman J., Setubal J.C.;
RT "The genome sequence of the plant pathogen *Xylella fastidiosa*,"
RT Nature 406:151-159(2000).
DR EMBL: AE003939; AAF84831.1; -.
DR HSSP: P23911; 1C8U.
DR InterPro: IPR003703; Acyl_CoA_thio.
DR Pfam: PF02551; Acyl_CoA_thio; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 310 AA; 35218 MW; CDF0DD4B24A064A CRC64;

Query Match 33.6%; Score 559; DB 16; Length 310;
Best Local Similarity 39.5%; Pred. No. 8.5e-48;
Matches 117; Conservative 54; Mismatches 107; Indels 18; Gaps 5;

QY 21 TVVNLPLEDEDLFRGRHHVWPAKRLFGGQIVGOALVAAKSVSDVHSHLCYFVRAG 80
DB 26 TWKLSLERLEDNLFGRQNDIGTKYVFGGQILAAQAQNTIENSRLHAYFLRTG 85
QY 81 DPKLPVLYQVTRTSGSFVSRSKAVQHKPLIFCOASFOQAQSPMOHQFSMPTVPPP 140
DB 86 NIQOPIIYNVTRTGKSFSLRVRTAIGHQGVIFCTASFOCEDG-AEQHTNPVAVPPP 144
QY 141 ELLDCETLIDQLRDPNLQKRYPLALNRIAAQEVPIEKVPNPSPLSQLQRMPEKQMF 200
145 EDIAPTPEL-----PPEALPRMPLQMTWNLNLSGQFEFRVPNPNTLDLQKHSFQHLW 198
QY 201 VRARGYIGEDMKHCCVAAVYSDYAFGLTALLPHQWQVHP-----WVSLDHSWMFH 254
DB 199 LRNNAPLGD-RIELHQILTYASDFQLLETAYS-----RHGINVHTPOVQMASLDHALWFH 253
QY 255 APPRADHNLVCEPESFWAGSGRLVHGRWQDGLAVTCAQEGVIRKVPQVSESK 310
DB 254 RPFIDDDMLLYTLESPTAGAGLARGQFFTRNGVLNIAQGLMRTNHTTFSK 309

RESULT 7
ID Q9KT42 PRELIMINARY; PRT: 286 AA.
AC Q9KT42;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE ACYL-COA THIOESTERASE II.
GN VC1063.
OS *Vibrio cholerae*.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; *Vibrio*.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN-EL TOR NI6961 / SEROTYPE O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Emolaeva M.D., Vamathevan J., Bass S., Olin H., Dragoi I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*,"
RL Nature 406:477-483(2000).
DR EMBL: AE004187; AAF94222.1; -.
DR HSSP: P23911; 1C8U.
DR TIGR: VC1063;
DR InterPro: IPR003703; Acyl_CoA_thio.
DR Pfam: PF02551; Acyl_CoA_thio; 1.
KW Complete proteome.
SQ SEQUENCE 286 AA; 32231 MW; 1FAAF066220D0B00 CRC64;

Query Match 32.6%; Score 543; DB 16; Length 286;
Best Local Similarity 41.0%; Pred. No. 3.1e-46;
Matches 118; Conservative 48; Mismatches 100; Indels 22; Gaps 7;

QY 22 TVVNLPLEDEDLFRGRHHVWPAKRLFGGQIVGOALVAAKSVSDVHSHLCYFVRAG 81
DB 10 SLQLKLEELGSLYRGAENLGLPOVGGQVGOALSAARYVESDRTVSHSFYLYPGD 69
QY 82 PKLPVLYQVTRTSGSFVSRSKAVQHKPLIFCOASFOQAQSPMOHQFSMPTVPPP 141
DB 70 KXPPIIYDVNLRDGRKSFSTRVRKAIONGRDIFVLTASYHGDAPG-FEHQKMPVVP 128
QY 142 ELLDCETLIDQLRDPNLQKRYPLALNRIAAQEVPIEKVPNP-SPLSOLQRMPEKQMF 200
DB 129 NFASESELAQ-----IAHFLPEKLOAFCEPIEMRPVTVINPLXP-EKAEKQYLM 181
QY 201 VRARGYIGEDMKHCCVAAVYSDYAFGLTALLPH-----QWQHKVHFVSLDHSWMF 253
DB 182 IRANGMDPD-DQLIHQVLLGYASDWGFLVTLHPHGSLSLMTPKFO-----VATIDHSIW 235
QY 254 HAPFRADHNLVCEPESFWAGSGRLVHGRWQDGLAVTCAQEGVIR 301
DB 236 HRPFKMDMLLYTLESPTAGAGLARGQFFTRNGVLNIAQGLMRTNHTTFSK 283

RESULT 8
ID Q9KZJ2 PRELIMINARY; PRT: 288 AA.
AC Q9KZJ2;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE ACYL-COA THIOESTERASE.
GN SCG8A.07.
OS *Streptomyces coelicolor*.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycetales; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RA Brown S.P., Harris D.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RA Bentley S.D., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RX MEDLINE=97000351; PubMed=8843436;

RA Redenbach M., Kieser H.M., Denapaita D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
DR EMBL: AL353863; CAB88937.1; -.
DR HSSP; P23911; 1C8U.
DR InterPro; IPR003703; Acyl_CoA_thio.
DR Pfam; PF02551; Acyl_CoA_thio; 1.
SQ SEQUENCE 288 AA; 32432 MW; 63E0F05AAE1DC5B4 CRC64;

Query Match 32.5%; Score 541.5; DB 2; Length 288;
Best Local Similarity 41.3%; Pred. No. 4.4e-46;
Matches 119; Conservative 49; Mismatches 105; Indels 15; Gaps 7;

QY 20 VTTVNLPLEDLPRGRHYWVPAKRLFGQIVGOALVAAAKSVSDVHSHCYEVRA 79
DB 8 LVALLDLQIEVNIIFGRSPESLQRFVGGVAGQALVAGRTTGDGRPVHSHAYFLRP 67
QY 80 GDPKLPVLYQVTRTGTSSFSVRSKAVQHGKPIFICQASFOQAQSPMOHQFSMPT--V 137
DB 68 GRPGVPIVQVDRDRGRSFTTRRTAVQGGRTIFLTASFHKPEPGSFHQPLP-PRKV 126
QY 138 PPEPELLDCETLIDQYLRDPNLQKRYPLALNRITAAQEVPIEIKVPNPSPLS--QLQRM 195
DB 127 PDPESLPTVADEVREHL-----GALPEQLERMARQ-PFDIRYDRLRWSAEVDEGAEP 179
QY 196 KQMFVVRARGYIGEGDMKHCVAAYISDYAFLTGALLPHQ--WQHKVHFVMSLDHSMWF 253
DB 180 RSNVWRAVGLGD-DPLVHTCALTVASDWTLLDVAIRIPVEPLNGPRGFMASLDHAWWF 238
QY 254 HAPFRADHWMLEYCESPWAGSGRLVHGRDQGLAVTCAQEGVIR 301
DB 239 HRFRADWFLYDQESPIATGGRLARGRIYDREGRMLSVVQEGFLR 286

RESULT 9
Q95Q68 ID Q95Q68 PRELIMINARY; PRT; 414 AA.

AC Q95Q68; 1
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE HYPOTHETICAL 46.2 KDA PROTEIN.
GN C37H5.13.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
NCBI_TaxID=6239;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE-99069613; PubMed-9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Davidson S., Gillam B.;
RT "The sequence of C. elegans cosmid C37H5.";
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; U88315; AAK68237.1; -.
KW Hypothetical protein.
SQ SEQUENCE 414 AA; 46244 MW; A13CB3AC514A26BF CRC64;

Query Match 31.1%; Score 518.5; DB 5; Length 414;
Best Local Similarity 37.0%; Pred. No. 1.4e-43;
Matches 121; Conservative 53; Mismatches 116; Indels 37; Gaps 8;

QY 14 DLRSVLVTVLNLEPLEDLPRGRHYWVPAKRL--FGGOIVGOALVAAAKSVSDVHSH 70
DB 95 DIRAGLIDTFLNLERVDKLNLYLARHLKGRNSLPVTVGGVIGQALSAATATVEGVFPN 154
QY 71 SLHCYFVRAGDPKLPVLYQVTRTGTSSFSVRSKAVQHGKPIFICQASFOQAQSPMOH 130
DB 155 SLHSFVQSGNVERPILYQVDRIRDGKSPCTRLVKALQDGEAIFTVOISFHRPADSIVH 214
QY 131 QFSMPTVPPEPELLDCETLIDQYLRDPNL-----OKRYPLALNRITAAQEVPIE 179
DB 215 QLPMEVPAPDSLEDLSDTFFERIKKNANIPPAALAMIFGKKEIPPAFFRI-----FSF 268
QY 180 KPVNPSPLSQLQRMPE-----KQMFVVRARGYIGEGDMKHCVAAYISDYAF 228
DB 269 RVPDIDSYLCLKDDTAGHGHPTDAYRSVWIKANENIGD-DPLRLHAAAAYISDATWI 327
QY 229 GTALLPHQHKVHFV-VSLDHSWMFHP-FRADHWMLEYCESPWAGSGRLVHGRWRQ 286
DB 328 ETALRPHSRKRGFTPSMALTLDSHSMHTDNFRVDDWMLYENHSTIAGGSRSLIEGKLWTR 387
QY 287 DGVLAVTCQAGEVIRV---KPVQVSESK 310
DB 388 DGRLVFTTQEALIRTAKNRPNASPAK 414

RESULT 10
Q9A2B6 ID Q9A2B6 PRELIMINARY; PRT; 284 AA.

AC Q9A2B6; 1
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE ACYL-COA THIOESTERASE II.
GN CC3649.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
OC Caulobacter.
NCBI_TaxID=69394;
RX [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 19089 / CB15;
RX MEDLINE-21173698; PubMed-11259647;
RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA Utterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
DR EMBL; AS006023; AAK25611.1; -.
DR HSSP; P23911; 1C8U.
DR TIGR; CC3649;
DR InterPro; IPR003703; Acyl_CoA_thio.
DR Pfam; PF02551; Acyl_CoA_thio; 1.
KW Complete proteome.
SQ SEQUENCE 284 AA; 31518 MW; F65A4C87C5560AF7 CRC64;

Query Match 30.1%; Score 501.5; DB 16; Length 284;
Best Local Similarity 41.5%; Pred. No. 4.4e-42;
Matches 118; Conservative 54; Mismatches 99; Indels 13; Gaps 9;

QY 23 VLNLEPLEDLPRGRHYWVPAKRLFGQIVGOALVAAAKSVSDVHSHCYEVRA 82
DB 10 ILDEPIEVNLFRCVSPNDGPRIFGGLVIAQALLAAAYKTPDRV-CHSLHAYFIRPDV 68

```

QY 83 KLPLVQVETRTGSSFSRVSKAVQHGKPIFICQAASFQAQSPMQHQFSMTPTVPDPPEE 142
Db      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 69 TAPVLVEVERARDGGTFTTRVAAIHQGEQIFNLAASTPQEDG-FHQSEMPASVDPS 127
Db      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 143 LLDCETLIDYOYLRDPLNQKRYPLALNRITAAQGEVPLEIKVPNP-SPLSQLORMPEKQMFV 201
Db      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 202 RARGYIGEGDMKMHCCVAAVYISDYAFGLTGLTALPHQ--WQHKVHFVMSLDHSMWFHAPFRA 259
Db      ||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 181 RAKAPLIGD-DVKMQAALAVASDMAFESALRPHGLIWTTPGIIQAASLDHAMWFHHPFNF 239
Db      ||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 260 DHMMLYECESPWAGSGRLVHGLRWRODGYLAVTCAQEGVTRVK 303
Db      ||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 240 NDWTLFAQDSPSASQGRGLYRGQMFSQDGKLLASVAQECMLRVR 283
Db      ||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

RESULT 11
087 PRELIMINARY; PRT; 320 AA.
Q9LO87;
Q9LO87;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE T1NG.10 PROTEIN.
GN T1NG.10.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eumetazoa II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Liu S., Chan A., Yu G., Lee J., Lenz C., Pham P., Sakano H.,
RA Toriumi M., Chin C., Chlou J., Choi E., Chung M., Gonzalez A.,
RA Hong B., Liu A., Vaysberg M., Altafi H., Brooks S., Buehler E.,
RA Chao Q., Conn L., Conway A., Hansen N., Johnson-Hopson C., Khan S.,
RA Kim C., Lam B., Miranda M., Nguyen M., Palm C., Shinn P.,
RA Southwick A., Davis R., Ecker J., Federspiel N., Theologis A.;
RT *The sequence of BAC T1NG from Arabidopsis thaliana chromosome 1.*;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC009273; AAF78401.1; -.
DR HSSP; P23911; 1C9U.
DR InterPro; IPR003703; Acyl_CoA_thio.
DR Pfam; PF02551; Acyl_CoA_thio; 1.
DR SEQUENCE 320 AA; 36157 MW; 03A82A96333E1BFC CRC64;

```

[illegible]

```

Qy 292 VTCAOEGVIR 301
Db 298 VSLTQALLR 307

RESULT 12
Q9CN69 PRELIMINARY; PRT; 292 AA.
AC Q9CN69;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE TSB.
GN TSB OR PM0570.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Pasteurella.
CN NCBI_TaxID=747;
RX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PM70;
RX MEDLINE=21145866; PubMed=11248100;
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RT *Complete genomic sequence of Pasteurella multocida Pm70.*;
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
RL EMBL: AE006092; AAK02654.1; -.
DR HSSP: P23911; 1C8U
DR InterPro: IPR003703; Acyl_CoA_thio.
DR Pfam: PF02551; Acyl_CoA_thio: 1.
KW Complete proteome.
SQ SEQUENCE 292 AA; 32977 MW; CB86741442225AF1 CRC64;

```

Query Match	27.8%	Score 463;	DB 16:	Length 292;
Best Local Similarity	38.8%;	Pred. No.	3.4e-38;	
Matches 114;	Conservative 48;	Mismatches 114;	Indels 18;	Gaps

Qy	23	VLNLEPDEDLFRGRHWVPKAKRLFGGOIVGOALVAAAKSVESDVHVHSLHCYFVRAGDP	82
	:	: :	
Dd	11	LQLERLDIDFLFRGKSEDIGLRQVEGGVVAQAUSAATQVAPPDRVLHSCHAYFLSPGDS	70
	:	: :	
Qy	83	KLPVLYQVERTRTGSSFSVRSVKAVOHGKPIFCOASFQOAQPSMQHQFSMPTVPPPEE	142
	:	: :	
Dd	71	OQHGIIVDVTLEGRNFTALRVRAIOHNEPICHTASFQKAEG-FDHQSMPDVESPDQ	129
	:	: :	
Qy	143	LLDCETLIDQYLNDPNLKRYPLNALNRIAAQEVPIEKPV---NPSPLSQLQMPEPKMF	199
	:	: :	
Dd	130	LI-AESAIIQ-----KLAPHIPEAIRAKFTAERPEDIRTKYNPNPFHCSQL----PAEQF	179
	:	: :	
Qy	200	-WVRARGYLEGDMKKHCCVAAVYISDYAFGLTALLPHQ--WQHKVFHFVSLDHSWMFHAP	256
	:	: :	
Dd	180	AWVKANGHAPO-DHKIQOCLLAYFSDFHCLLTALHPHQKGLOPGMKVATIHDHSIWFHRP	238
	:	: :	
Qy	257	FRAWHMLYCEBSPAGWGSRGLLVGHRLWRDGVLAVTCAGQGVTRVKPQVSSEK	310
	:	: :	
Dd	239	FDLDWLLYAIESNNFAARGLARGQIFDRQGLRIATTQQEGLRITYYPEKGEK	292
	:	: :	

RESULT 13	
QC321	
ID	PRELIMINARY; PRT; 368 AA.
OC	QC321;
DT	01-JUN-2001 (TrEMBLrel. 17, Created)
DT	01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT	01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE	ACVL-COA THIOESTERASE.
GN	TES1.
OS	Cochliobolus heterostrophus (Drechslera maydis).
OC	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Dothideomycetes;
OC	Pleosporales; Pleosporaceae; Cochliobolus.
OX	NCBI_TaxID=5016;
RN	[1]

[illegible]

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 19, 2002, 13:18:29 ; Search time 30.4 Seconds
(without alignments)
1136.314 Million cell updates/sec

Title: US-09-766-366-1
Perfect score: 1666
Sequence: 1 MGRAVATAALPGDLRSVLV.....VTCAOEGVIRVQVSESKL 311

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues
number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802.*
1: /SIDS1/gcgdata/hold-genseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/hold-genseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/hold-genseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/hold-genseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/hold-genseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/hold-genseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/hold-genseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/hold-genseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/hold-genseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/hold-genseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/hold-genseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/hold-genseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/hold-genseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/hold-genseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/hold-genseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/hold-genseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/hold-genseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/hold-genseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/hold-genseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/hold-genseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/hold-genseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/hold-genseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1666	100.0	311	19	AAW58530 Human peroxisomal
2	1666	100.0	311	20	AAV08604 Human PXT protein
3	1666	100.0	311	22	AAV077835 Protein sequence o
4	1666	100.0	311	22	AAV073420 Human peroxisomal
5	1284	77.1	331	22	ABG20249 Novel human diagno
6	628.5	37.7	270	22	AAU04004 Thioesterase type
7	553	33.2	286	20	AAV08619 E. coli TEII prote
8	489	29.4	271	21	AAV24648 Arabidopsis thalia
9	489	29.4	377	21	AAV24647 Arabidopsis thalia
10	489	29.4	385	21	AAV24646 Arabidopsis thalia
11	484.5	29.1	106	22	ABG20246 Novel human diagno

12	466.5	28.0	282	22	AAV01582 C glutamicum prote
13	466.5	28.0	282	22	AAV06590 Corynebacterium gl
14	466.5	28.0	282	22	AAV06591 Corynebacterium gl
15	449.5	27.0	368	22	AAU03392 Cochliobolus heter
16	433	26.0	298	22	AAU02058 Propionibacterium
17	426	25.6	104	22	ABG20245 Novel human diagno
18	273	16.4	349	20	AAV08620 Yeast CoA thioeste
19	201	12.1	81	22	ABG20244 Novel human diagno
20	196	11.8	140	22	AAU05800 Propionibacterium
21	180.5	10.8	420	22	ABG25236 Novel human diagno
22	145	8.7	91	22	ABG08886 Lipid degradation
23	145	8.7	91	22	ABG08336 P patens lipid met
24	127	7.6	43	22	ABG20247 Novel human diagno
25	111.5	6.7	557	22	ABG36202 Novel human diagno
26	109.5	6.6	591	22	AAV39282 Human polypeptide
27	109.5	6.6	594	22	AAV01068 Human polypeptide
28	100	6.0	1580	18	AAV23719 Platenolide syntha
29	100	6.0	1580	18	AAV22609 Drosophila sp. mus
30	91	5.5	1425	22	ABG65506 Drosophila sp. mus
31	91	5.5	1425	22	AAE04383 Human HuiPRG-3 pro
32	90.5	5.4	1234	22	AAE07171 Human secreted pro
33	90.5	5.4	1268	22	ABB11796 Acetyl CoA carboxy
34	89	5.3	2240	15	AAV67819 Maize acetyl CoA c
35	89	5.3	2325	17	AAU05590 Maize ACCase enzym
36	89	5.3	2325	19	AAV44687 Maize acetyl CoA c
37	89	5.3	2325	21	AAU03388 Cochliobolus heter
38	88.5	5.3	1743	22	AAU28244 Novel human secret
39	88	5.3	278	22	ABV58608 Drosophila melanog
40	86	5.2	1816	22	ABV08747 Drosophila melanog
41	85.5	5.1	1070	18	ABV70969 Drosophila melanog
42	85	5.1	361	22	ABV69572 Transcriptional ac
43	85	5.1	746	22	AAU02193 Enterococcus faeca
44	84	5.0	648	22	
45	83.5	5.0	331	22	

ALIGNMENTS

RESULT 1
AAW58530
ID AAW58530 standard; Protein; 311 AA.

AC AAW58530;

XX 01-SEP-1998 (first entry)

DT Human peroxisomal thioesterase.

XX Human; peroxisomal thioesterase; PXT; diagnosis; inflammatory disease;

KW fatty acid metabolism disorder; cancer; drug screening.

XX Homo sapiens.

OS US5776753-A.

PN 07-JUL-1998.

PD 11-JUN-1997; 97US-0872784.

PF 11-JUN-1997; 97US-0872784.

PR (INCY-) INCYTE PHARM INC.

PA Corley NC, Hillman JL, Shah P;

XX WPI: 1998-398039/34.

DR N-PSDB; AAV31109.

XX New nucleic acid encoding human peroxisomal thioesterase - useful

PT for diagnosis, treatment and prevention of fatty acid metabolism

PT disorders, cancer and inflammatory disease, and in drug screening

XX

PS Claim 9; Fig 1; 30pp; English.

XX The present sequence represents human peroxisomal thioesterase (PxTE).
CC The cDNA encoding PxTE was isolated from a human BRAIN09 clone 2150905
CC cDNA library. PxTE and its agonists are used in treatment, diagnosis
CC and prevention of disorders of fatty acid metabolism e.g. neuronal
CC disorders, Zellweger or Down's syndrome, Refsum's disease, epilepsy,
CC Alzheimers disease, depression, tardive dyskinesia, multiple sclerosis,
CC Parkinson's disease, schizophrenia and Tourette's syndrome. Fragments of
CC PxTE DNA can also be used for diagnosis and monitoring (in hybridisation
CC or amplification methods), including determining a predisposition to
CC disease, for gene mapping and to detect PxTE-encoding nucleic acid,
CC particularly after amplification. PxTE can be used to raise antibodies
CC (Ab) used therapeutically, in competitive drug screens, as immunocassay
CC reagents for diagnosis and monitoring of diseases and for purification of
CC natural PxTE, and to screen for specific binding agents antagonists e.g.
CC specific Ab, antisense sequences or ribozymes which are used in treatment
CC of cancers and inflammatory diseases (e.g. adult respiratory distress
CC syndrome, asthma, diabetes mellitus, microbial and other infections).
CC Administration of proteins, agonist or antagonists, optionally combined
CC with conventional pharmaceuticals providing a synergistic effect, is
CC e.g. orally, by injection or topically.

XX Sequence 311 AA;

Query Match 100.0%; Score 1666; DB 19; Length 311;
Best Local Similarity 100.0%; Pred. No. 4.5e-177;
Matches 311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGRAVATAALPPGDLRSVLTVTLNLEPLEDLFRGRHYWVPKRLFGGQIVGQALVAAA 60

Db 1 mgravataalppgdlrsvlvtvtnlepledlfrgrhywvpakrlfggqivgqalvaaa 60

Qy 61 KVSSEDEVHSHLCYFVRAGDKPLVLYOYVTRTSGSSFSRSVKAVOHGKPIFICOASF 120

Db 61 kvssedevhshlcyfvragdplvlyoyvtrtsgssfsrsvkavghgkpficqasf 120

Qy 121 QQAQSPMQHQSMTPTVPPPEELLDCTETLIDQYLDRPNLQKRYPLALNRIAAAQVPIEIK 180

Db 121 qqaqspmqhqsmtptvpppeelldcetlidqylrdpnlkryplalnriaaaqevpieik 180

Qy 181 PVNPSPLSQLQRMPEKQMFVWRARGYIGEGDMKHCVCVAAAYISDYAFGLTALLPHQWQHK 240

Db 181 pvnpsplsqlqrmepkqmfvrargyigegdmkmcvcvaaayisdyaflgtallphqwgkh 240

Qy 241 VHFVMSLDHSMWFHAPFRADHWMLYECESPWAGSGRLVHGRLWRQDGLAVTCAQEGVI 300

Db 241 vhfvmvslhsmwfhapfradhwmlyecespwagsgrglvhgrlwrqdgglavtcaqegvi 300

Qy 301 RVKPOVSESGL 311

Db 301 rvkpvqseskl 311

RESULT 2

AA08604

ID AA08604 standard; Protein; 311 AA.

XX AA08604;

XX 05-AUG-1999 (first entry)

XX Human PxTE protein.

XX PxTE; peroxisomal thioesterase; treatment; disorder; biosynthesis;
KW fatty acid metabolism; chain-termination; fatty acid; cancer; deficiency;
KW inflammation; adrenoleukodystrophy Refsum's disease; mental retardation;
KW disordered neuronal migration; hypotonia; tapetoretinal degeneration;
KW sensorineural hearing loss; cystic change; kidney; skeletal change.

XX Homo sapiens.

XX

US5911984-A.

PD 15-JUN-1999.

XX 19-JUN-1998; 98US-0100851.

XX 11-JUN-1997; 97US-0872784.

PR 19-JUN-1998; 98US-0100851.

XX (INCY-) INCYTE PHARM INC.

XX Corley NC, Hillman JL, Shah P;

PI WPI; 1999-357169/30.

DR N-PSDB; AAX77452.

XX Peroxisomal thioesterase enzymes useful for treating disorders
PT associated with fatty acid metabolism

PS Claim 1; Fig 1A-C; 29pp; English.

XX This invention describes a novel human peroxisomal thioesterase enzyme
CC (PxTE) and its use to treat disorders associated with fatty acid
CC metabolism. Thioesterases catalyze the chain-terminating step in the de
CC novo biosynthesis of fatty acids. PxTE may be used to treat cancer, the
CC inflammation and disorders associated with fatty acid metabolism and the
CC expression of PxTE, such as adrenoleukodystrophy Refsum's disease, and
CC acid deficiency. Patients suffering from defective peroxisomal fatty
CC acid metabolism exhibit disordered neuronal migration, hypotonia, mental
CC retardation, tapetoretinal degeneration, sensorineural hearing loss,
CC cystic changes in the kidney, skeletal changes and death.

XX Sequence 311 AA;

Query Match 100.0%; Score 1666; DB 20; Length 311;
Best Local Similarity 100.0%; Pred. No. 4.5e-177;
Matches 311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGRAVATAALPPGDLRSVLTVTLNLEPLEDLFRGRHYWVPKRLFGGQIVGQALVAAA 60

Db 1 mgravataalppgdlrsvlvtvtnlepledlfrgrhywvpakrlfggqivgqalvaaa 60

Qy 61 KVSSEDEVHSHLCYFVRAGDKPLVLYOYVTRTSGSSFSRSVKAVOHGKPIFICOASF 120

Db 61 kvssedevhshlcyfvragdplvlyoyvtrtsgssfsrsvkavghgkpficqasf 120

Qy 121 QQAQSPMQHQSMTPTVPPPEELLDCTETLIDQYLDRPNLQKRYPLALNRIAAAQVPIEIK 180

Db 121 qqaqspmqhqsmtptvpppeelldcetlidqylrdpnlkryplalnriaaaqevpieik 180

Qy 181 PVNPSPLSQLQRMPEKQMFVWRARGYIGEGDMKHCVCVAAAYISDYAFGLTALLPHQWQHK 240

Db 181 pvnpsplsqlqrmepkqmfvrargyigegdmkmcvcvaaayisdyaflgtallphqwgkh 240

Qy 241 VHFVMSLDHSMWFHAPFRADHWMLYECESPWAGSGRLVHGRLWRQDGLAVTCAQEGVI 300

Db 241 vhfvmvslhsmwfhapfradhwmlyecespwagsgrglvhgrlwrqdgglavtcaqegvi 300

Qy 301 RVKPOVSESGL 311

Db 301 rvkpvqseskl 311

RESULT 3

AAAG77835

ID AAG77835 standard; protein; 311 AA.

XX AAG77835;

XX 03-JAN-2002 (first entry)

XX Protein sequence of human peroxisomal thioesterase (PxTE).

XX Human peroxisomal thioesterase; PxTE; immunogenic; cytostatic;
KW neuroleptic; antiinflammatory; neuroprotective;
KW cerebroprotective; Down's syndrome; gene therapy; cancer; inflammation;
KW fatty acid metabolism disorders; neuronal disorders; Alzheimer's disease;
KW amnesia; Huntington's disease; cerebropathorenal syndrome;
KW hybridisation probe; adrenomyeloneuropathy; schizophrenia;
KW adrenoleukodystrophy; Crohn's disease; adenocarcinoma; leukaemia;
KW Addison's disease; adult respiratory distress syndrome.
XX Homo sapiens.
XX Key Location/Qualifiers
FH Region 309..311
FT /note= "C-terminal SKL (ser-lys-leu) matrix targeting
FT sequence"
XX US01024786-A1.
XX 27-SEP-2001.
XX 18-JAN-2001; 2001US-0766366.
XX 11-JUN-1997; 97US-0872784.
PR 19-JUN-1998; 98US-0100851.
PR 09-MAR-1999; 99US-0265294.
XX (INCY-) INCYTE PHARM INC.
XX Hillman JL, Shah P, Corley NC;
XX WPI: 2001-638498/73.
DR N-PSDB; AAF78821.
XX New human peroxisomal thioesterase polypeptide and polynucleotide,
PT useful in diagnosing, preventing and treating cancer, inflammation and
PT disorders associated with fatty acid metabolism -
XX Claim 1; Fig 1; 33pp; English.
XX The protein sequence represents the human peroxisomal thioesterase
CC enzyme (PxTE) of the invention. The invention also includes the
CC polynucleotide sequence coding for this protein. Also the invention
CC provides expression vectors, host cells, agonists, antibodies and
CC antagonists. The invention also provides methods for treating disorders
CC associated with expression of PxTE. The invention has the following
CC activities of being cytostatic, antiinflammatory, neuroprotective,
CC neuroleptic, cerebroprotective and neuroleptic although no supporting data
CC is given. The PxTE polypeptides are useful in diagnosing, preventing and
CC treating cancer (e.g. adenocarcinoma or leukaemia), inflammation (e.g.
CC colon and small intestine afflicted by Crohn's disease) and disorders
CC associated with fatty acid metabolism, such as neuronal disorders
CC including Alzheimer's disease, amnesia, Down syndrome, Huntington's
CC disease, or schizophrenia by using the PxTE polypeptides in gene therapy.
CC In one embodiment, an antagonist of PxTE may be administered to a subject
CC to prevent or treat inflammation caused by conditions such as Addison's
CC disease or adult respiratory distress syndrome. The PxTE polynucleotides
CC can be used to detect and quantify gene expression in biopsied tissue. In
CC the diagnosis of conditions, disorders or diseases associated with PxTE
CC expression e.g. adrenoleukodystrophy, adrenomyeloneuropathy, or
CC cerebropathorenal syndrome, to detect activation or induction of various
CC cancers, to generate hybridization probes useful for mapping naturally
CC occurring genomic sequences, and for screening libraries in various drug
CC screening techniques.
SQ Sequence 311 AA;
Query Match 100.0%; Score 1666; DB 22; Length 311;
Best Local Similarity 100.0%; Pred. No. 4.5e-177;
Matches 311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGRAVATAALPPGDLRSVLVTTVNLNLEPDLDFRGRHYWVPAKRLFGGQIVGQALVAAA 60

Db 1 mgravataaalppgdlrsvlvttvnllepldedlfrgrhywvpakrlfggqivgqalvaa 60
QY 61 KYSVEDVHVHSLHCYFYVRAGDPKLPVLYQVETRTGSSFSVRSKAVOHGKPTIFCOASF 120
Db 61 kysvedvhvshlhcylfyragdpklpvlyqvverttgssfsvrskavohgkptifcoasf 120
QY 121 QQAQSPMQHQFSMPTVPPPEELDCEFTLIDQYLRDPLQKRYPLALNRIAAQEVPIEK 180
Db 121 qqaqspmqhqfsmptvpppeelldcetlldqylrdpnlkryplalnriaaqevpiek 180
QY 181 PVNPSLSQLQRMPEPKQMFVWRARGYIGEGDMKHCVCVAAAYISDYAFGLCTALLPHQWQHK 240
Db 181 pvnpslsqqlrmepkqmfvwrargyigegdmkmcvcvaaaysdyafglctallphqwk 240
QY 241 VHEMVSJDHSMWFHAPFRADHMLYECESPWAGSGRGLVHGRLWRQDGLVAVTCAQEGVI 300
Db 241 vhemvsldhsmwfhapfradhmlyecespwagsgrglvhgrlwrqdgvlavtcaqegvi 300
QY 301 RVKPPQVSESKL 311
Db 301 rvkppqvseskl 311
RESULT 4
ID AAB73420 standard; Protein: 311 AA.
XX AAB73420;
XX 02-JUL-2001 (first entry)
DE Human peroxisomal thioesterase, PxTE.
KW Human; peroxisomal thioesterase; PxTE; fatty acid biosynthesis;
KW chain termination; detection; diagnosis; monitoring; hybridisation;
KW inflammation; cancer; defective fatty acid metabolism; neuronal disorder;
KW adrenoleukodystrophy; adrenomyeloneuropathy; cerebropathorenal syndrome;
KW Zellweger syndrome; Refsum's disease; Alzheimer's disease; amnesia;
KW amyotrophic lateral sclerosis; bipolar disorder; Huntington's disease;
KW Down's syndrome; multiple sclerosis; neurofibromatosis; schizophrenia;
KW Parkinson's disease; paranoid psychosis; Tourette's syndrome;
KW drug screening.
XX Homo sapiens.
XX Key Location/Qualifiers
FH Region 309..311
FT /note= "Peroxisomal matrix targeting sequence"
FT US6210890-B1.
XX 03-APR-2001.
XX 09-MAR-1999; 99US-0265294.
XX 11-JUN-1997; 97US-0872784.
PR 19-JUN-1998; 98US-0100851.
XX (INCY-) INCYTE GENOMICS INC.
XX Hillman JL, Shah P, Corley NC;
XX WPI: 2001-289629/30.
DR N-PSDB; AAF76400.
XX Detecting human peroxisomal thioesterase (PxTE) in a biological sample,
PT useful for diagnosing disorders associated with fatty acid metabolism,
PT comprises hybridizing a PxTE polynucleotide complement with the sample
PT nucleic acid -
XX Claim 1; Fig 1A-C; 30pp; English.
XX

CC This sequence represents a human peroxisomal thioesterase, PXT_E.
CC PXT_E, like other thioesterases, catalyses the chain-terminating
CC step in the de novo biosynthesis of fatty acids. PXT_E is expressed
CC in neuronal, gastrointestinal and secretory tissues, and cells associated
CC with inflammation and cancer. The invention relates to a method of
CC detecting nucleic acids encoding PXT_E involving the hybridisation of a
CC PXT_E nucleic acid probe to PXT_E nucleic acids in a sample, and the
CC subsequent detection of the hybridisation complex. The method is useful
CC for the diagnosis of cancer, inflammation and disorders associated with
CC fatty acid metabolism, and may also be used to monitor the treatment of
CC a patient with a disorder associated with PXT_E expression. Disorders
CC associated with defective fatty acid metabolism include neuronal
CC disorders such as adrenoleukodystrophy, adrenomyeloneuropathy,
CC cerebrotendinous degeneration (Zellweger syndrome), Refsum's disease,
CC Alzheimer's disease, amnesia, amyotrophic lateral sclerosis, bipolar
CC disorder, Huntington's disease, Down's syndrome, multiple sclerosis,
CC neurofibromatosis, Parkinson's disease, paraneoplastic psychosis,
CC schizophrenia, or Tourette's syndrome. In addition to their use as
CC probes in the method of the invention, PXT_E nucleic acids may also
CC be used in chromosomal mapping, detecting polymorphisms associated
CC with disease, in gene therapy, and in drug screening techniques. The
CC PXT_E protein may also be used in drug screening, and to raise
CC antibodies.

XX Sequence 311 AA;
SQ
Query Match 100.0%; Score 1666; DB 22; Length 311;
Best Local Similarity 100.0%; Pred. No. 4.5e-177;
Matches 311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGRAVATAALPGDLSRVLTVTNLNLEPDLDFRGRHWVPVAKRLFGGIVGQALVAAA 60
Db 1 mgravataalpgdlsrvltvttnlepldedlfrgrhwvpvapakrlfggivgqalvaa 60
Qy 61 KSVSEDVHSHLCYFVFRAGDPKLPVLYQVTRTSGSFVSRSVKAQVHGKPIFCOASF 120
Db 61 ksvsedvvhshlchcyfvragdplkplvyqverttsgsfvsrvskavqhgkpficqasf 120
Qy 121 QQAQSPMQHQSMTPTVPPPEELDCETLIDQYLRDPNLQKRYPLALNRITAAQEVPIEK 180
Db 121 qqaqspmqhqsmtptvpppeelddcetlidqylrdpnlqkryplalnrlaaevppei 180
Qy 181 PVNPSPLSOLRMEPKQMFVWRARGYIGEGDMKMHCCVAAVYISDYAFGLTALLPHQOHK 240
Db 181 pvnpsplsolrmepekmtvwrargyigegdmkmmhccvvaayisdyafigtallphqwhk 240
Qy 241 VHFVMSLDHSMWFFHAPFRADHWMLYECESPWAGGSRGLVHGRLWRQDGVILAVTCAQEGVI 300
Db 241 vhfmsldhsmwffhapfradhwmllyecespwaggsrglvhgrlwrqdgvlavtcaqegvi 300
Qy 301 RVKPOVSESKL 311
Db 301 rvkpvqvsesk1 311

RESULT 5
ABG20249
ID ABG20249 standard; Protein; 331 AA.
XX
XX AC ABG20249;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #20240.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
FN WO200175067-A2.
XX

PD 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
XX
XX MPI; 2001-639362/73.
XX N-PSDB; AAS84436.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity -
XX
XX Claim 20; SEQ ID No 50608; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX and gene mapping, and in recombinant production of (II). The
XX polynucleotides are also used in diagnostics as expressed sequence tags
XX for identifying expressed genes. (I) is useful in gene therapy techniques
XX to restore normal activity of (II) or to treat disease states involving
XX (II). (II) is useful for generating antibodies against it, detecting or
XX quantitating a polypeptide in tissue, as molecular weight markers and as
XX a food supplement. (II) and its binding partners are useful in medical
XX imaging of sites expressing (II). (I) and (II) are useful for treating
XX disorders involving aberrant protein expression or biological activity.
XX The polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. ABG00010-ABG30377 represent novel human
XX diagnostic amino acid sequences of the invention.
XX Note: The sequence data for this patent did not appear in the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 331 AA;
Query Match 77.1%; Score 1284; DB 22; Length 331;
Best Local Similarity 81.7%; Pred. No. 2e-134;
Matches 263; Conservative 11; Mismatches 36; Indels 12; Gaps 12;
Qy 2 GRAVATAALPGDLSRVLTVTNLNLEPDLDFRGRHWVPVAKRLFGGIVGQALVAAA 61
Db 10 gqgcgdrgappgdlsrvltvttnlepldedlfrgrhwvpvapakrlfggivgqanvaat 69
Qy 62 SVSEDEVHSHLCYFVFRAGDPKLPVLYQVTRTSGSFVSRSVKAQVHGKPIFCO 117
Db 70 svsedvvhshlchcyfvragdplkplvyqverttsgsfvsrvskavqhgkpficq 129
Qy 118 ASFQQAQSPMQHQSMTPTVPPPEELDCETLIDQYLRDPNLQKRYPLALNRITAAQEV 175
Db 130 asfqaqspmqhqsmtptlptteelxvtrtldqylrdpnlqkryplalnrlaaev 189
Qy 176 PIEIKPVNPSPL-SQLQRMPEKQMFVWRARGYIGEG-DMKMHCCVAAVYISDYAFGLTALL 233
Db 190 pieikpvnpvpsplrpaeeepkqmfvwrargyigegtkstpwkmhccvvaayisdyafigtala 249
Qy 234 P-HQW-OHKVHFH-VSLDHS-WMFHAPFRADHWMLYECESPWAGGSRGLVHGRLWRQDGV 289
Db 250 altqwpstrchfhgppldhnpwfhapfradhwmllyecespwaggsrglvhgrlwrqdg 309
Qy 290 LAVTCAQEGVIRVKPOVSESKL 311
Db 310 lavtcaqegvirvkpvqvsesk1 331

PA (CORR) CORNELL RES FOUND INC.
 XX Yoder OC, Turgeon BC, Lu S;
 XX WPI; 2001-367672/38.
 XX New isolated nucleic acid molecule from a plant pathogen useful in
 PT preventing plant pathogenic infections -
 XX Disclosure; Fig 22; 132pp; English.
 XX The sequence represents a thioesterase type II consensus sequence. These
 CC thioesterases include Cochliobolus heterostrophus TES1, thioesterase I,
 CC the gene for which is located in the same genomic region as the peptide
 CC synthetase, CPS1. CPS1 is an enzyme thought to be involved in the
 CC production of peptide toxins, which are involved in the pathogenic
 CC infection of corn crops. The nucleic acids and proteins can be used as
 CC targets for anti-fungal compounds to prevent fungal corn infection and
 CC the nucleic acids can be used in gene therapy to alter the biosynthetic
 CC pathway for the peptide toxins to lower the pathogenicity of the fungi.
 XX Sequence 270 AA:
 Query Match 37.7%; Score 628.5; DB 22; Length 270;
 Best Local Similarity 44.2%; Pred. No. 2.1e-61;
 Matches 138; Conservative 11; Mismatches 78; Indels 85; Gaps 4;
 QY 25 NLEPLEDELFGRGHYVWP---AKRLFGGOIVGOALVAAAKSVSEDDVHSLHCYFVRAGD 81
 Db :||||:||||| :|:||||:||||| :|:||||:||||| :|:||||:||||| :|:||||:|||||
 4 dlepleexlfrgxxxexxgarrxfggqvvggalvaaxktvdedxxvhslhcyflragd 63
 QY 82 PKLPVLYQVTRTGSSFSVRSVKAYQHGKPIFCQASFOQAQSPMQHQFSMPTVPPPE 141
 Db :||||:||||| :|:||||:||||| :|:||||:||||| :|:||||:||||| :|:||||:|||||
 64 skxpxlyxverrxgxfstrrxvavghgkpfixmtasfxxqxxqxxqxxmpxap--- 120
 QY 142 ELLDCEETLIDQLRDPNLQKRYPLALNRITAAQAEVPIEIKPVNPSLSQLOMQMPEKQFWV 201
 Db :||||:||||| :|:||||:||||| :|:||||:||||| :|:||||:||||| :|:||||:|||||
 121 -----xxxxxepxqgvwx 133
 QY 202 RARGYIGEDMKMHCCVAAAYISDYAFGLTALLPQWQHVKHF----- 243
 Db :||||:||||| :|:||||:||||| :|:||||:||||| :|:||||:||||| :|:||||:|||||
 134 targaaxpdxxxxhxcalaymsdxxlflgtalqphxixxxxxxxxxxxxxxxxxxxxxxxx 193
 QY 244 -----MVSLDHSWMFH--APFRADHMLYECESPWAGSGRLVHGLRWQRDGV 289
 Db :||||:||||| :|:||||:||||| :|:||||:||||| :|:||||:||||| :|:||||:|||||
 194 xxxdxxxxxxqmasldhsmwfhrrxxpfradewllyexespsaggggglvrgexytdqgv 253
 290 LAVTCAQEGVIR 301
 Db :|:||||| :|:||||:||||| :|:||||:||||| :|:||||:||||| :|:||||:|||||
 254 lvascvqegvxx 265
 RESULT 7
 AAY08619
 ID AAY08619 standard; Protein; 286 AA.
 XX AC AAY08619;
 XX 05-AUG-1999 (first entry)
 XX E. coli TEII protein.
 XX PXT; peroxisomal thioesterase; treatment; disorder; biosynthesis;
 KW fatty acid metabolism; chain-termination; fatty acid; cancer; deficiency;
 KW inflammation; adrenoleukodystrophy Refsum's disease; mental retardation;
 KW disordered neuronal migration; hypotonia; tapetoretinal degeneration;
 KW sensorineural hearing loss; cystic change; kidney; skeletal change; TEII.
 XX Escherichia coli.
 OS US5911384-A.
 XX PN
 XX

PD 15-JUN-1999.
 XX 19-JUN-1998; 98US-0100851.
 XX 11-JUN-1997; 97US-0872784.
 PR 19-JUN-1998; 98US-0100851.
 XX (INCY-) INCYTE PHARM INC.
 XX Corley NC, Hillman JL, Shah P;
 XX WPI; 1999-357169/30.
 DR Peroxisomal thioesterase enzymes useful for treating disorders
 PT associated with fatty acid metabolism
 PS Disclosure; Fig 2A-B; 29pp; English.
 XX This invention describes a novel human peroxisomal thioesterase enzyme
 CC (PXT) and its use to treat disorders associated with fatty acid
 CC metabolism. Thioesterases catalyze the chain-terminating step in the de
 CC novo biosynthesis of fatty acids. PXT may be used to treat cancer,
 CC inflammation and disorders associated with fatty acid metabolism and the
 CC expression of PXT, such as adrenoleukodystrophy Refsum's disease, and
 CC PXT deficiency. Patients suffering from defective peroxisomal fatty
 CC acid metabolism exhibit disordered neuronal migration, hypotonia, mental
 CC retardation, tapetoretinal degeneration, sensorineural hearing loss,
 CC cystic changes in the kidney, skeletal changes and death. This sequence
 CC represents the Escherichia coli TEII protein which is used to illustrate
 CC the method of the invention
 XX SQ Sequence 286 AA;
 Query Match 33.2%; Score 553; DB 20; Length 286;
 Best Local Similarity 42.1%; Pred. No. 6e-53;
 Matches 125; Conservative 53; Mismatches 91; Indels 28; Gaps 10;
 QY 15 LRSVLVTTVLNLEPLDELFGRGHYVWPYAKRLFGGOIVGOALVAAAKSVSEDDVHSLHC 74
 Db :||||:||||| :|:||||:||||| :|:||||:||||| :|:||||:||||| :|:||||:|||||
 5 lxnll--tllnlekieeglfgrgsedlgrqvfggqvgaalyaaketvpeerlvhsfhs 62
 QY 75 YFVRAGDPKLPVLYQVTRTGSSFSVRSVKAVOHGKPIFCQASFOQAQSPMQHQFSM 134
 Db :||||:||||| :|:||||:||||| :|:||||:||||| :|:||||:||||| :|:||||:|||||
 63 yfrrpdsxkpllydvettldrgnsfsarrvaaigngkpflymtasf-qapeagfheqkcm 121
 QY 135 PTVPPEELLDCEETLIDQ---YLRDPLNQLKRYPLALNRITAAQAEVPIEIKPVN-PSPLSQ 190
 Db :|:||||| :|:||||:||||| :|:||||:||||| :|:||||:||||| :|:||||:|||||
 122 psapapdg-lpsetqiaqlahlppvldkf-----icdrplevrpvefhaplkq- 171
 QY 191 QRMEPKQMFVWRARGYIGEDMKMHCCVAAAYISDYAFGLTALLPQWQHVKHF----- 244
 Db :|:||||| :|:||||:||||| :|:||||:||||| :|:||||:||||| :|:||||:|||||
 172 hvaephrrqvwirangsvpd-dlrhqvilygsdlnflpvalqp---hgflpeglq1 226
 QY 245 VSLDHSWMFHAPFRADHMLYECESPWAGSGRLVHGLRWQRDGVLAIVTCAQEGVIR 301
 Db :|:||||| :|:||||:||||| :|:||||:||||| :|:||||:||||| :|:||||:|||||
 227 atidhsmwfhrrpnlnewllyvestaasargfvrgfytqdgvlvavstvgvmmr 283
 RESULT 8
 AAG24648
 ID AAG24648 standard; Protein; 271 AA.
 XX AC AAG24648;
 XX 17-OCT-2000 (first entry)
 XX Arabidopsis thaliana protein fragment SEQ ID NO: 28404.
 XX Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.

OS Arabidopsis thaliana.
XX EP1033405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 25-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127452.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151308.
PR 07-SEP-1999; 99US-0151930.
PR 10-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.

PR	15-SEP-1999;	99US-0154018.	
PR	16-SEP-1999;	99US-0154039.	
PR	20-SEP-1999;	99US-0154779.	
PR	22-SEP-1999;	99US-0155139.	
PR	23-SEP-1999;	99US-0155486.	
PR	24-SEP-1999;	99US-0155659.	
PR	28-SEP-1999;	99US-0156458.	
PR	29-SEP-1999;	99US-0156596.	
PR	04-OCT-1999;	99US-0157117.	
PR	05-OCT-1999;	99US-0157753.	
PR	06-OCT-1999;	99US-0157865.	
PR	07-OCT-1999;	99US-0158029.	
PR	08-OCT-1999;	99US-0158232.	
PR	12-OCT-1999;	99US-0158369.	
PR	13-OCT-1999;	99US-0159293.	
PR	13-OCT-1999;	99US-0159294.	
PR	13-OCT-1999;	99US-0159295.	
PR	14-OCT-1999;	99US-0159329.	
PR	14-OCT-1999;	99US-0159330.	
PR	14-OCT-1999;	99US-0159331.	
PR	14-OCT-1999;	99US-0159637.	
PR	14-OCT-1999;	99US-0159638.	
PR	18-OCT-1999;	99US-0159584.	
PR	21-OCT-1999;	99US-0160741.	
PR	21-OCT-1999;	99US-0160767.	
PR	21-OCT-1999;	99US-0160768.	
PR	21-OCT-1999;	99US-0160770.	
PR	21-OCT-1999;	99US-0160814.	
PR	21-OCT-1999;	99US-0160815.	
PR	22-OCT-1999;	99US-0160980.	
PR	22-OCT-1999;	99US-0160981.	
PR	22-OCT-1999;	99US-0160989.	
PR	25-OCT-1999;	99US-0161404.	
PR	25-OCT-1999;	99US-0161405.	
PR	25-OCT-1999;	99US-0161406.	
PR	26-OCT-1999;	99US-0161359.	
PR	26-OCT-1999;	99US-0161360.	
PR	26-OCT-1999;	99US-0161361.	
PR	28-OCT-1999;	99US-0161920.	
PR	28-OCT-1999;	99US-0161992.	
PR	28-OCT-1999;	99US-0161993.	
PR	29-OCT-1999;	99US-0162142.	
Query Match 29.4%; Score 489; DB 21; Length 271;			
Best Local Similarity 38.8%; Pred. No. 7.7e-46;			
Matches 101; Conservative 52; Mismatches 93; Indels 14; Gaps 4;			
14	DLRSVLVTTLNLEPLDELRGRHWVPK-----RLFGGOIVGOALVAAAKSVSEDVH	68	
17	dkerclvedillylepldlnvryg---ftppnptygkyvggqlvgqalaaasktvetmki	73	
69	VHSLHCYFVRAGDPKLPVLQYVERTGSSFSVRSYKAVQHGKPIFTCOASFQOAPSPM	128	
74	vhnthcyllvgdlnpiiydvnrldgnnfatrsvdarqkgkttlftasfqkkgqgfi	133	
129	QHQSMPVPPEELLDCTLDQYLRDNLKRYPLALNRIAAQ---EVPDIEKPVNPS	185	
134	hgestmpitpapelilpreemlerivteplldry---rnqvateisvpfpdirfcpn	190	
186	PLSOLQRMPEKQMFVWRARGYIGEGDMKMHCCVAAVYISDYAFUGTALLPHQWQHKVFMV	245	
191	rstqknksprikywfrakgkiddqalhrwcwafasdlifatislnphrrsgmsvaai	250	
246	SLDHSWMFAPFRADHWMLY	265	
251	slahsmwfrhprvradwlif	270	
RESULT 9			
AAAG24647			
ID	AAG24647 standard; Protein; 377 AA.		
XX			
AC	AAG24647;		

XX	17-OCT-2000 (first entry)	
DT	Arabidopsis thaliana protein fragment SEQ ID NO: 28403.	
XX	Arabidopsis thaliana.	
DE	Protein identification; signal transduction pathway; metabolic pathway;	
XX	hybridisation assay; genetic mapping; gene expression control; promoter;	
KW	termination sequence.	
KW	Arabidopsis thaliana.	
OS	Arabidopsis thaliana.	
XX	EP1033405-A2.	
PN	06-SEP-2000.	
XX	25-FEB-2000; 2000EP-0301439.	
PD	25-FEB-1999; 99US-0121825.	
PF	05-MAR-1999; 99US-0123180.	
XX	09-MAR-1999; 99US-0123548.	
PR	23-MAR-1999; 99US-0125788.	
PR	25-MAR-1999; 99US-0126264.	
PR	29-MAR-1999; 99US-0126785.	
PR	01-APR-1999; 99US-0127462.	
PR	06-APR-1999; 99US-0128234.	
PR	08-APR-1999; 99US-0128714.	
PR	16-APR-1999; 99US-0129845.	
PR	19-APR-1999; 99US-0130077.	
PR	21-APR-1999; 99US-0130449.	
PR	23-APR-1999; 99US-0130510.	
PR	23-APR-1999; 99US-0130891.	
PR	28-APR-1999; 99US-0131449.	
PR	30-APR-1999; 99US-0132048.	
PR	04-MAY-1999; 99US-0132407.	
PR	05-MAY-1999; 99US-0132484.	
PR	06-MAY-1999; 99US-0132485.	
PR	06-MAY-1999; 99US-0132486.	
PR	07-MAY-1999; 99US-0132487.	
PR	11-MAY-1999; 99US-0132863.	
PR	14-MAY-1999; 99US-0134256.	
PR	14-MAY-1999; 99US-0134218.	
PR	14-MAY-1999; 99US-0134219.	
PR	14-MAY-1999; 99US-0134221.	
PR	14-MAY-1999; 99US-0134370.	
PR	18-MAY-1999; 99US-0134768.	
PR	19-MAY-1999; 99US-0134941.	
PR	20-MAY-1999; 99US-0135124.	
PR	21-MAY-1999; 99US-0135353.	
PR	24-MAY-1999; 99US-0135629.	
PR	25-MAY-1999; 99US-0136021.	
PR	27-MAY-1999; 99US-0136392.	
PR	28-MAY-1999; 99US-0136782.	
PR	01-JUN-1999; 99US-0137222.	
PR	03-JUN-1999; 99US-0137528.	
PR	04-JUN-1999; 99US-0137502.	
PR	07-JUN-1999; 99US-0137724.	
PR	08-JUN-1999; 99US-0138094.	
PR	10-JUN-1999; 99US-0138540.	
PR	10-JUN-1999; 99US-0138847.	
PR	14-JUN-1999; 99US-0139119.	
PR	16-JUN-1999; 99US-0139452.	
PR	16-JUN-1999; 99US-0139453.	
PR	17-JUN-1999; 99US-0139492.	
PR	18-JUN-1999; 99US-0139454.	
PR	18-JUN-1999; 99US-0139455.	
PR	18-JUN-1999; 99US-0139456.	
PR	18-JUN-1999; 99US-0139457.	
PR	18-JUN-1999; 99US-0139458.	
PR	18-JUN-1999; 99US-0139459.	
PR	18-JUN-1999; 99US-0139460.	
PR	18-JUN-1999; 99US-0139461.	
PR	18-JUN-1999; 99US-0139462.	
PR	18-JUN-1999; 99US-0139463.	


```
||||| | | | | | |
Db 357 slhsmwfhrpvradwllf 376

RESULT 10
AAG24646
ID AAG24646 standard; Protein; 385 AA.
XX
AC AAG24646;
XX
XX 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 28402.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
EP1033405-A2.
PD
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 28-APR-1999; 99US-0130891.
XX 28-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 30-APR-1999; 99US-0132407.
XX 04-MAY-1999; 99US-0132484.
XX 05-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 07-MAY-1999; 99US-0132487.
XX 07-MAY-1999; 99US-0132863.
XX 11-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
XX 14-MAY-1999; 99US-0134370.
XX 18-MAY-1999; 99US-0134768.
XX 19-MAY-1999; 99US-0134941.
XX 20-MAY-1999; 99US-0135124.
XX 21-MAY-1999; 99US-0135353.
XX 24-MAY-1999; 99US-0135629.
XX 25-MAY-1999; 99US-0136021.
XX 27-MAY-1999; 99US-0136392.
XX 28-MAY-1999; 99US-0136782.
XX 01-JUN-1999; 99US-0137222.
XX 03-JUN-1999; 99US-0137528.
XX 04-JUN-1999; 99US-0137502.
XX 07-JUN-1999; 99US-0137724.
XX 08-JUN-1999; 99US-0138094.
XX 10-JUN-1999; 99US-0138540.
XX 10-JUN-1999; 99US-0138847.
XX 14-JUN-1999; 99US-0139119.
XX 16-JUN-1999; 99US-0139452.
XX 16-JUN-1999; 99US-0139453.
XX 17-JUN-1999; 99US-0139492.
XX 18-JUN-1999; 99US-0139454.
XX 18-JUN-1999; 99US-0139455.
XX 18-JUN-1999; 99US-0139456.
XX 18-JUN-1999; 99US-0139457.
XX 18-JUN-1999; 99US-0139458.
XX 18-JUN-1999; 99US-0139459.
XX 18-JUN-1999; 99US-0139460.
XX 18-JUN-1999; 99US-0139461.
XX 18-JUN-1999; 99US-0139462.
XX 18-JUN-1999; 99US-0139463.
XX 18-JUN-1999; 99US-0139750.
XX 18-JUN-1999; 99US-0139763.
XX 21-JUN-1999; 99US-0139817.
XX 22-JUN-1999; 99US-0139899.
XX 23-JUN-1999; 99US-0140353.
XX 23-JUN-1999; 99US-0140354.
XX 24-JUN-1999; 99US-0140695.
XX 28-JUN-1999; 99US-0140823.
XX 29-JUN-1999; 99US-0140991.
XX 30-JUN-1999; 99US-0141287.
XX 01-JUL-1999; 99US-0141842.
XX 02-JUL-1999; 99US-0142154.
XX 06-JUL-1999; 99US-0142055.
XX 08-JUL-1999; 99US-0142390.
XX 08-JUL-1999; 99US-0142803.
XX 09-JUL-1999; 99US-0142920.
XX 12-JUL-1999; 99US-0142977.
XX 13-JUL-1999; 99US-0143542.
XX 14-JUL-1999; 99US-0143624.
XX 15-JUL-1999; 99US-0144005.
XX 16-JUL-1999; 99US-0144085.
XX 16-JUL-1999; 99US-0144086.
XX 19-JUL-1999; 99US-0144325.
XX 19-JUL-1999; 99US-0144331.
XX 19-JUL-1999; 99US-0144332.
XX 19-JUL-1999; 99US-0144333.
XX 19-JUL-1999; 99US-0144334.
XX 19-JUL-1999; 99US-0144335.
XX 20-JUL-1999; 99US-0144352.
XX 20-JUL-1999; 99US-0144632.
XX 20-JUL-1999; 99US-0144884.
XX 21-JUL-1999; 99US-0144814.
XX 21-JUL-1999; 99US-0145086.
XX 21-JUL-1999; 99US-0145088.
XX 22-JUL-1999; 99US-0145085.
XX 22-JUL-1999; 99US-0145087.
XX 22-JUL-1999; 99US-0145089.
XX 22-JUL-1999; 99US-0145192.
XX 23-JUL-1999; 99US-0145145.
XX 23-JUL-1999; 99US-0145218.
XX 23-JUL-1999; 99US-0145224.
XX 26-JUL-1999; 99US-0145276.
XX 27-JUL-1999; 99US-0145913.
XX 27-JUL-1999; 99US-0145918.
XX 27-JUL-1999; 99US-0145919.
XX 28-JUL-1999; 99US-0145951.
XX 02-AUG-1999; 99US-0146386.
XX 02-AUG-1999; 99US-0146388.
XX 02-AUG-1999; 99US-0146389.
XX 03-AUG-1999; 99US-0147038.
XX 04-AUG-1999; 99US-0147204.
XX 04-AUG-1999; 99US-0147302.
XX 05-AUG-1999; 99US-0147192.
XX 05-AUG-1999; 99US-0147260.
XX 06-AUG-1999; 99US-0147303.
XX 06-AUG-1999; 99US-0147416.
XX 09-AUG-1999; 99US-0147493.
XX 09-AUG-1999; 99US-0147935.
XX 10-AUG-1999; 99US-0148171.
XX 11-AUG-1999; 99US-0148319.
XX 12-AUG-1999; 99US-0148341.
XX 13-AUG-1999; 99US-0148565.
XX 13-AUG-1999; 99US-0148684.
XX 16-AUG-1999; 99US-0149368.
```



```

PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 29.4%; Score 489; DB 21; Length 385;
Best Local Similarity 38.8%; Pred. No. 1.3e-45;
Matches 101; Conservative 52; Mismatches 93; Indels 14; Gaps 4;

QY 14 DLRSVLVTTLNLEPDEDLFRGRHYWVPAK-----RLFGQIYGOALVAAAKSVSEVDH 68
DB 131 dkerclvedilylepladlnvryg---ftppnptygkyvgqlyvgqalaaskvtetnki 187

QY 69 VHSLSHCYFVRAGDPKPLVLYQVETRTGSSFSVRSKAVQHGKPIFICQASFQQAQPSPM 128
DB 188 vhnfcyfillvgdinpilydvnlrdgnnfatrsvdarqkgkiftlfasfqkqgfi 247

```

```

QY 129 QHQFSMTVPDPPELDCETLIDQYLRDNPLOKRYPLALNRTAAQ---EVPTEIKPVNPS 185
DB 248 hgestmhtpaptellpreemlerivtepllpdy---rnqvateisvfpidircepn 304
QY 186 PLSOLQRMEPKOMFWVRARVIGEGDKMKHCCVAAAYISDYAFGLGTALLPHOMQHKVHFV 245
DB 305 rstqknspprlkywfrakgklstdqalhrccvafasdlifatislnphrrregmsvaal 364
QY 246 SLDRSMWFHAPFRADHNMPLY 265
DB 365 sldhsmwfhprvradddwlf 384

RESULT 11
ABG20246
ID ABG20246 standard; Protein; 106 AA.
XX
AC ABG20246;
XX
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #20237.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WC200175067-A2.
PD 11-OCT-2001.
PF 30-MAR-2001; 2001WO-US08631.
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
WPI: 2001-639362/73.
N-PSDB; AAS84433.
XX
New isolated polynucleotide and encoded polypeptides, useful in
diagnostics, forensics, gene mapping, identification of mutations
responsible for genetic disorders or other traits and to assess
biodiversity -
Claim 20; SEQ ID No 50605; 103pp; English.
The invention relates to isolated polynucleotide (I) and
polypeptide (II) sequences. (I) is useful as hybridisation probes,
polymerase chain reaction (PCR) primers, oligomers, and for chromosome
and gene mapping, and in recombinant production of (II). The
polynucleotides are also used in diagnostics as expressed sequence tags
for identifying expressed genes. (I) is useful in gene therapy techniques
to restore normal activity of (II) or to treat disease states involving
(II). (II) is useful for generating antibodies against it, detecting or
quantitating a polypeptide in tissue, as molecular weight markers and as
a food supplement. (II) and its binding partners are useful in medical
imaging of sites expressing (II). (I) and (II) are useful for treating
disorders involving aberrant protein expression or biological activity.
The polypeptide and polynucleotide sequences have applications in
diagnostics, forensics, gene mapping, identification of mutations
responsible for genetic disorders or other traits to assess biodiversity
and to produce other types of data and products dependent on DNA and
amino acid sequences. ABG0010-ABG30377 represent novel human
diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
specification, but was obtained in electronic format directly from WIPO
at ftp.wipo.int/pub/published_pct_sequences.
XX

```


PR 14-JUL-1999; 99DE-1033006.
PR 27-AUG-1999; 99DE-1040764.
PR 27-AUG-1999; 99DE-1040765.
PR 27-AUG-1999; 99DE-1040766.
PR 27-AUG-1999; 99DE-1040830.
PR 27-AUG-1999; 99DE-1040831.
PR 27-AUG-1999; 99DE-1040832.
PR 27-AUG-1999; 99DE-1040833.
PR 31-AUG-1999; 99DE-1041378.
PR 31-AUG-1999; 99DE-1041379.
PR 03-SEP-1999; 99DE-1041395.
PR 03-SEP-1999; 99DE-1042077.
PR 03-SEP-1999; 99DE-1042078.
PR 03-SEP-1999; 99DE-1042079.
PR 03-SEP-1999; 99DE-1042088.
XX PA
XX (BADI) BASF AG.

PI Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;
XX

XX WPI; 2001-071486/08.
XX N-PSDB; AAF67823.

XX Corynebacterium glutamicum nucleic acids encoding membrane construction
PT and membrane transport proteins or their portions, useful for typing or
PT identifying C. glutamicum or related bacteria, and as markers for
PT transformation -

XX Claim 20; Page 392-393; 1119pp; English.

XX AAF67743 to AAF68080 encode the Corynebacterium glutamicum membrane
CC construction and membrane transport (MCT) proteins given in AAB76510 to
CC AAB76847. The MCT nucleic acids and proteins are useful in the
CC identification of microorganisms which can be used to produce fine
CC chemicals, for modulating fine chemical production in C. glutamicum or
CC related bacteria (e.g. Brevibacterium lactofermentum), the typing or
CC identification of C. glutamicum or related bacteria, as reference points
CC for mapping C. glutamicum genome, and as markers for transformation.
CC AAF68082 and AAF68082 represent sequencing primers which are used in an
CC example from the present invention.

XX Sequence 282 AA;

Query Match 28.0%; Score 466.5; DB 22; Length 282;
Best Local Similarity 36.6%; Pred. No. 2.6e-43;
Matches 107; Conservative 45; Mismatches 119; Indels 21; Gaps 7;

OY 20 VTTVLNLEPLDEDLFGRHYWVPKAKLFGGOIVGOALVAAKSVSEDEVHSLHCYFVRA 79

D 4 iedilleidridiypviesylartlfgggvaqaivaactvdkaftvshlghyfiap 63

OY 80 GPKPLVLYOVERTRTGSSFSVRSKAVQHGKPTFCQASQQAQSPMOHQSMTYVPP 139

Db 64 gptapailvdvrdgksyvtvrigdqvifsmqasfhrqdeg-iehmdkmrkypa 122

OY 140 PEELLDCETLIDQYLRDPNLOKRYPLALNRIAAQ--EVPTEIKPVNPSPLSQLORMKPKQ 197

Db 123 pdei-----kgtvempisrrvldewaewdirvpqdgqlsdfate--q 167

OY 198 MFVVRARGVIGEDMKMCCVAAVYISDYAFGLTALLPHQOHKHYFMVSLDHSMMFHAFPE 257

Db 168 awvircatldp-nptfhqcslytsmdtlhslavhpqge-kmq-masldhavwflrpf 224

OY 258 RADHWMLYECESPWAGSGRLVHGRWRQDQVLAVTCAQEGVIRVKFQVSES 309

Db 225 rvdewillydqrspassgralthrlnqgqdlvaivnqegmtrtlbegaqs 276

RESULT 14

AAB76591

ID AAB76591 standard; Protein; 282 AA.

XX

AC AAB76591;

XX 11-APR-2001 (first entry)

XX Corynebacterium glutamicum MCT protein SEQ ID NO:164.

XX Corynebacterium glutamicum; Brevibacterium lactofermentum; MCT;
KW membrane construction and membrane transport protein; petroleum spill;
KW hydrocarbon degradation; gram positive aerobic bacterium; marker;
KW identification; microorganism; fine chemical production; transformation;
KW genome mapping; genetic engineering.

XX Corynebacterium glutamicum.

XX WO200100805-A2.

XX 04-JAN-2001.

XX 23-JUN-2000; 2000WO-IB00926.

XX 25-JUN-1999; 99US-0141031.

PR 08-JUL-1999; 99DE-1031454.

PR 08-JUL-1999; 99DE-1031478.

PR 08-JUL-1999; 99DE-1031563.

PR 09-JUL-1999; 99DE-1032122.

PR 09-JUL-1999; 99DE-1032124.

PR 09-JUL-1999; 99DE-1032125.

PR 09-JUL-1999; 99DE-1032128.

PR 09-JUL-1999; 99DE-1032180.

PR 09-JUL-1999; 99DE-1032182.

PR 09-JUL-1999; 99DE-1032190.

PR 09-JUL-1999; 99DE-1032191.

PR 09-JUL-1999; 99DE-1032209.

PR 09-JUL-1999; 99DE-1032212.

PR 09-JUL-1999; 99DE-1032227.

PR 09-JUL-1999; 99DE-1032228.

PR 09-JUL-1999; 99DE-1032229.

PR 09-JUL-1999; 99DE-1032230.

PR 14-JUL-1999; 99DE-1032927.

PR 14-JUL-1999; 99DE-1033005.

PR 14-JUL-1999; 99DE-1033006.

PR 27-AUG-1999; 99DE-1040764.

PR 27-AUG-1999; 99DE-1040765.

PR 27-AUG-1999; 99DE-1040766.

PR 27-AUG-1999; 99DE-1040830.

PR 27-AUG-1999; 99DE-1040831.

PR 27-AUG-1999; 99DE-1040832.

PR 27-AUG-1999; 99DE-1040833.

PR 31-AUG-1999; 99DE-1041378.

PR 31-AUG-1999; 99DE-1041379.

PR 31-AUG-1999; 99DE-1041395.

PR 03-SEP-1999; 99DE-1042077.

PR 03-SEP-1999; 99DE-1042078.

PR 03-SEP-1999; 99DE-1042079.

PR 03-SEP-1999; 99DE-1042088.

XX (BADI) BASF AG.

XX Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;

XX WPI; 2001-071486/08.

XX N-PSDB; AAF67824.

XX Corynebacterium glutamicum nucleic acids encoding membrane construction
PT and membrane transport proteins or their portions, useful for typing or
PT identifying C. glutamicum or related bacteria, and as markers for
PT transformation -

XX Claim 20; Page 395-396; 1119pp; English.

XX

XX AAF67743 to AAF68080 encode the Corynebacterium glutamicum membrane

CC construction and membrane transport (MCT) proteins given in AAB76510 to

CC AAB76847. The MCT nucleic acids and proteins are useful in the

CC identification of microorganisms which can be used to produce fine
CC chemicals, for modulating fine chemical production in C. glutamicum or
CC related bacteria (e.g. Brevibacterium lactofermentum), the typing or
CC identification of C. glutamicum or related bacteria, as reference points
CC for mapping C. glutamicum genome, and as markers for transformation.
CC AAF68082 and AAF68082 represent sequencing primers which are used in an
CC example from the present invention.

XX SQ Sequence 282 AA;

Query Match 28.0%; Score 466.5; DB 22; Length 282;
Best Local Similarity 36.6%; Pred. No. 2.6e-43;
Matches 107; Conservative 45; Mismatches 119; Indels 21; Gaps 7;

Qy 20 VTTVLNLEPLDLDLFRGRHYWPAKRLFGCGQIVGQALVAAAKSVSDVHVSLSLHCYFVRA 79
Db 4 Iedilteieidridyrgviesylartfggvaqaalvaathvdkafthslhyfiap 63
80 GDPKLPVLYQVVERTRTGSSFSVRSKAVQHGKPIFTICQASFQQAQSPMQHQFSMPRTVPP 139
Db 64 gdpapalyldrvrdgksyvrsvrglqdgveifsmqasfhrqdeg-iehmckmrkvp 122
Qy 140 PEEILDCELTLDYLRDNPMLQKRYPLALNRIAAQ--EVPLEIKPVNPSPLSQLQRMPEKQ 197
Db 123 pdei-----kgtvermpissrrvidewaeewdlrvipqqlelsdfate-q 167
Qy 198 MFVWRARGYIGSGDMKMCVAAAYISDYAFGLTALLPHQWQHVHPMVSLDHSMMWFHAPF 257
Db 168 awwircatldp-nptfhqcsitylsdmtllhsalvphqge-kmq-masidhavwflrpf 224
Qy 258 RADHWMLEYCESPWAGGSRGLVHGLWRQDGVLAVTCAQEGVIRVXPQVSES 309
Db 225 rvdewillydqrspssasgralthrlnfdqgdlvaivnqegmtrtlhegaqs 276

RESULT 15

AAU03392 ID AAU03392 standard; Protein; 368 AA.

XX AC AAU03392;

XX DT 12-SEP-2001 (first entry)

XX DE Cochliobolus heterostrophus TES1, thioesterase 1.

XX KW CPS1; peptide synthetase; peptide toxin; fungal pathogen;
corn crop infection; TES1; thioesterase.

XX Cochlilobolus heterostrophus.

XX FH Key Location/Qualifiers
XX FT Active-site 72..74

XX PN WO200138489-A2.

XX PD 31-MAY-2001.

XX XX 22-NOV-2000; 2000WO-US32227.

XX XX 23-NOV-1999; 99US-0448215.

XX PA (CORR) CORNELL RES FOUND INC.

XX XX Yoder OC, Turgeon BC, Lu S;

XX XX WPI; 2001-367672/38.

XX DR N-PSDB; AAS05896.

XX PT New isolated nucleic acid molecule from a plant pathogen useful in

XX preventing plant pathogenic infections -

XX XX Example 4; Fig 18; 132pp; English.

XX
CC
CC
CC
CC
CC
CC
CC
CC
CC
XX
SQ

The sequence represents Cochliobolus heterostrophus TES1, thioesterase 1,
the gene for which is located in the same genomic region as the peptide
synthetase, CPS1. CPS1 is an enzyme thought to be involved in the
production of peptide toxins, which are involved in the pathogenic
infection of corn crops. The nucleic acids and proteins can be used as
targets for anti-fungal compounds to prevent fungal corn infection and
the nucleic acids can be used in gene therapy to alter the biosynthetic
pathway for the peptide toxins to lower the pathogenicity of the fungi.

Sequence 368 AA;

Query Match 27.0%; Score 449.5; DB 22; Length 368;
Best Local Similarity 32.6%; Pred. No. 3.1e-41;
Matches 117; Conservative 44; Mismatches 107; Indels 91; Gaps 11;

Qy 11 PGDLRSVLVTTLNLEPL--DEDLFRG-RHYWVP--AKRLFGQIVGQALVAAAKSVS 64
Db 8 ppadhsqsaienveitqtadldpnlftnrplwhppgargifggaiaqltsaaqktvd 67
Qy 65 EDVHVHSLHCYFVRAGDPKLPVLYQVVERTRTGSSFSVRSKAVQHGKPIFTICQASF--Q 121
Db 68 pdftvshmbcyfilagnseipviyhvervrsqkafartvqargrqnvtfttmsfvrqn 127
Qy 122 QAQPSMQHQFSMPRTVPPPEELDCETLIDQYLRDNPMLQKRYPLALNRIAAQEVPIEKP 181
Db 128 sggaqkvehlypmpdvapkegsd-----dlktpnqg-----spfqtlplienad 174
Qy 182 VNPSPLSQLQRMPEKQMFVWRARGYIG-EGDMKMHCCVAAAYISDYAFVIGTA-----LLPH 235
Db 175 dsdkphtkkcrq-----wikargkispagghheahlsalaysmsdsyfigtvarahklly 228
Qy 236 QWQHK-----wikargkispagghheahlsalaysmsdsyfigtvarahklly 240
Db 229 snqrksarssidedvkllemdaelqrqsfvnesdkqrirleirkaedlaksqdakpe 288
Qy 241 VHFVMSLDHSMWFHAP--FRADHWMLEYCESPWAGGSRGLVHGLWRQDGVLAVTCAQE 297
Db 289 igmmvslldhtlyfhnprsfadewiftemetpwagdgrglvsqrmytkdgtllascvqe 347

Search completed: August 19, 2002, 13:19:29
Job time: 60 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 19, 2002, 13:18:29 ; Search time 13.12 Seconds
(without alignments)
578.991 Million cell updates/sec

Title: US-09-766-366-1
Perfect score: 1666
Sequence: 1 MGRAVATAALPPGDLRSVLV.....VTCAQGVIRVKPOVSESKL 311

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/prodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/prodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/prodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/prodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/prodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/prodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1666	100.0	311	1	US-08-872-784-1
2	1666	100.0	311	2	US-09-100-851-1
3	1666	100.0	311	4	US-09-265-294-1
4	553	33.2	286	1	US-08-872-784-3
5	553	33.2	286	2	US-09-100-851-3
6	553	33.2	286	4	US-09-265-294-3
7	273	16.4	349	1	US-08-872-784-4
8	273	16.4	349	2	US-09-100-851-4
9	273	16.4	349	4	US-09-265-294-4
10	100	6.0	1580	2	US-08-804-227C-11
11	100	6.0	1580	2	US-08-804-198-5
12	89	5.3	2325	3	US-08-417-089-6
13	89	5.3	2325	4	US-08-695-651-6
14	89	5.3	2325	4	US-08-930-285-6
15	89	5.3	2325	4	US-08-695-421-6
16	83.5	5.0	529	5	PCT-US92-00282-7
17	80.5	4.8	516	4	US-08-867-611-6
18	80.5	4.8	516	5	PCT-US92-06965A-11
19	80.5	4.8	798	4	US-08-867-611-36
20	79.5	4.8	1068	4	US-09-085-199B-11
21	79.5	4.7	412	4	US-09-461-697-77
22	78.5	4.7	525	1	US-08-160-861-4
23	78.5	4.7	680	1	US-08-542-363-4
24	78.5	4.7	680	4	US-09-100-089-4
25	78.5	4.7	1350	2	US-08-319-866-9
26	78	4.7	481	3	US-08-787-091-2
27	77.5	4.7	732	1	US-08-317-522A-5

Sequence 5, Appli
Sequence 5, Appli
Sequence 5, Appli
Sequence 5, Appli
Sequence 5, Appli
Sequence 24, Appli
Sequence 24, Appli
Sequence 4, Appli
Sequence 9, Appli
Sequence 31, Appli
Sequence 26, Appli
Sequence 4, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 22, Appli
Sequence 22, Appli
Sequence 9, Appli
Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-08-872-784-1
; Sequence 1, Application US/08872784
; Patent No. 5776753
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Shah, Purvi
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN PEROXISOMAL THIOESTERASE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA: US/08/872,784
; APPLICATION NUMBER: US/08/872,784
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0293 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 311 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRAINOT09
; CLONE: 2150905
US-08-872-784-1

Query Match 100.0%; Score 1666; DB 1; Length 311;
Best Local Similarity 100.0%; Pred. No. 1.5e-182;
Matches 311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGRAVATAALPPGDLRSVLVTVTLNLEPLDEDLFRGRHWWPAKRLFGQIIVGOALVAAA 60
Db 1 MGRAVATAALPPGDLRSVLVTVTLNLEPLDEDLFRGRHWWPAKRLFGQIIVGOALVAAA 60
QY 61 KVSSEDVHSHLHCYFVRAGDPKLPVLYQVTRTGSFVSRSKAVQHGKPIFICQASF 120
Db 61 KVSSEDVHSHLHCYFVRAGDPKLPVLYQVTRTGSFVSRSKAVQHGKPIFICQASF 120
QY 121 QQAQPSMQHQSMTPTVPPPELDCETLIDQYLRDPLNLRKRYPLALNRIAAQEVPIEK 180
Db 121 QQAQPSMQHQSMTPTVPPPELDCETLIDQYLRDPLNLRKRYPLALNRIAAQEVPIEK 180
QY 181 PVNPSLSQLORMPEKQMFWRARGYIGEGDMKMHCCVAAAYISDYAFGLTALLPHQWQHK 240
Db 181 PVNPSLSQLORMPEKQMFWRARGYIGEGDMKMHCCVAAAYISDYAFGLTALLPHQWQHK 240
QY 241 VHFVMSLDHSMWFHAPFRADHMLYECESPWAGSGRGLVHGRWRQDGLVAVTCAQEGVI 300
Db 241 VHFVMSLDHSMWFHAPFRADHMLYECESPWAGSGRGLVHGRWRQDGLVAVTCAQEGVI 300
301 RVKPOVSESCL 311
301 RVKPOVSESCL 311

RESULT 2
US-09-100-851-1
; Sequence 1, Application US/09100851
; Patent No. 5911984
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: HUMAN PEROXISOMAL THIOESTERASE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,851
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/872,784
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0293 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 311 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRAINOT09
; CLONE: 2150905
US-09-100-851-1

Query Match 100.0%; Score 1666; DB 2: Length 311;
Best Local Similarity 100.0%; Pred. No. 1.5e-182; Indels 0; Gaps 0;
Matches 311; Conservative 0; Mismatches 0;
QY 1 MGRAVATAALPPGDLRSVLVTVTLNLEPLDEDLFRGRHWWPAKRLFGQIIVGOALVAAA 60
Db 1 MGRAVATAALPPGDLRSVLVTVTLNLEPLDEDLFRGRHWWPAKRLFGQIIVGOALVAAA 60
QY 61 KVSSEDVHSHLHCYFVRAGDPKLPVLYQVTRTGSFVSRSKAVQHGKPIFICQASF 120
Db 61 KVSSEDVHSHLHCYFVRAGDPKLPVLYQVTRTGSFVSRSKAVQHGKPIFICQASF 120
QY 121 QQAQPSMQHQSMTPTVPPPELDCETLIDQYLRDPLNLRKRYPLALNRIAAQEVPIEK 180
Db 121 QQAQPSMQHQSMTPTVPPPELDCETLIDQYLRDPLNLRKRYPLALNRIAAQEVPIEK 180
QY 181 PVNPSLSQLORMPEKQMFWRARGYIGEGDMKMHCCVAAAYISDYAFGLTALLPHQWQHK 240
Db 181 PVNPSLSQLORMPEKQMFWRARGYIGEGDMKMHCCVAAAYISDYAFGLTALLPHQWQHK 240
QY 241 VHFVMSLDHSMWFHAPFRADHMLYECESPWAGSGRGLVHGRWRQDGLVAVTCAQEGVI 300
Db 241 VHFVMSLDHSMWFHAPFRADHMLYECESPWAGSGRGLVHGRWRQDGLVAVTCAQEGVI 300
301 RVKPOVSESCL 311
301 RVKPOVSESCL 311
RESULT 3
US-09-265-294-1
; Sequence 1, Application US/09265294
; Patent No. 6210890
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Shah, Purvi
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN PEROXISOMAL THIOESTERASE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/265,294
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,851
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0293 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 311 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRAINOT09
; CLONE: 2150905

US-09-265-294-1

Query Match 100.0%; Score 1666; DB 4; Length 311;
Best Local Similarity 100.0%; Pred. No. 1.5e-182;
Matches 311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGRAVATAALPPGDLRSVLTTLNLEPLDEDLFRGRHYWVPAKRLFGGQIVGQALVAAA 60
DB 1 MGRAVATAALPPGDLRSVLTTLNLEPLDEDLFRGRHYWVPAKRLFGGQIVGQALVAAA 60
QY 61 KVSSEDVHVHSLCYFVRAGDPKLPVLYQVETRTGSSFSVRSVKAVQHGKPIFCQASF 120
DB 61 KVSSEDVHVHSLCYFVRAGDPKLPVLYQVETRTGSSFSVRSVKAVQHGKPIFCQASF 120
QY 121 QQAQSPMQHQSMTPTVPPPEELDCETLIDQYLRDNLQKRYPLALNRNIAAEVPIEIK 180
DB 121 QQAQSPMQHQSMTPTVPPPEELDCETLIDQYLRDNLQKRYPLALNRNIAAEVPIEIK 180
QY 181 PVNPSLSQLQRMPEKQMFVVRARGYIGEDMKMHCCVAAIYSDYAFGLTALLPHQWQHK 240
DB 181 PVNPSLSQLQRMPEKQMFVVRARGYIGEDMKMHCCVAAIYSDYAFGLTALLPHQWQHK 240
QY 241 VHFVYSLDHSNMFHAPRADHMLYECESPWAGSGRLVHGRLWRQDGLVAVTCAQEGVI 300
DB 241 VHFVYSLDHSNMFHAPRADHMLYECESPWAGSGRLVHGRLWRQDGLVAVTCAQEGVI 300
QY 301 RVKQVSESKL 311
DB 301 RVKQVSESKL 311

RESULT 4

US-08-872-784-3
; Sequence 3, Application US/08872784
; Patent No. 5776753
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Shah, Purvi
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN PEROXISOMAL THIOESTERASE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/872,784
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0293 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 286 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 147932
US-08-872-784-3

Query Match 33.2%; Score 553; DB 1; Length 286;
Best Local Similarity 42.1%; Pred. No. 4.2e-55;
Matches 125; Conservative 53; Mismatches 91; Indels 28; Gaps 10;
QY 15 LRSVLVTTLNLEPLDEDLFRGRHYWVPAKRLFGGQIVGQALVAAKSYSEDVHVHSLHC 74
DB 5 LKNLL--TLNLEKTEEGLFQGSDELGLRQVGGVQALYAAKETVPEERLVHSEHS 62
QY 75 YFVRAGDPKLPVLYQVETRTGSSFSVRSVKAVQHGKPIFCQASFQQAQSPMQHQSFM 134
DB 63 YFLRPGDSKKPIIYDVTLRDGNFSARRVAAIQNGKPIFYMTASP-OAPEAGFEHOKTM 121
QY 135 PTVPPPEELDCETLIDQ--YLRDNLQKRYPLALNRNIAAEVPIEIKPVN-PSPLSOL 190
DB 122 PSAPAPDG-LPSETQIAQSLAHLPLPVLDKDF-----ICDRPLEVRPVEFHNPLKG- 171
QY 191 QRMPEKQMFVVRARGYIGEDMKMHCCVAAIYSDYAFGLTALLPHQWQHKVHFM----- 244
DB 172 HVAEPHROWIRANGSVDP-DLRVHQYLLGYASDLNPLVALQP----HGIGFLEPGIQI 226
QY 245 VSLDSHMFHAPRADHMLYECESPWAGSGRLVHGRLWRQDGLVAVTCAQEGVIR 301
DB 227 ATIDHSMFHRPFLNLEWLLYSVESTSASSARGFVRGEFTQDQVLVAVTQEGVMR 283

RESULT 5

US-09-100-851-3
; Sequence 3, Application US/09100851
; Patent No. 5911984
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Shah, Purvi
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN PEROXISOMAL THIOESTERASE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,851
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/872,784
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0293 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 286 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:

LIBRARY: GenBank
CLONE: 147932
US-09-100-851-3

Query Match 33.28; Score 553; DB 2; Length 286;
Best Local Similarity 42.18; Pred. No. 4.2e-55;
Matches 125; Conservative 53; Mismatches 91; Indels 28; Gaps 10;

QY 15 LRSVLVTVLNLEPLDEDLFRGRHYWVPAKRLFGGQIVGOALVAAKSVSDVHVHSLHC 74
Db 5 LKLL--TLNLEKIEGLFRGQSEDLGRVFGGVQVGOALYAARETVPEERLVHSFHS 62
QY 75 YFVRAGDPKLPVLYQVTRTGSFVSRSVKAQVHGKPIFCQAFQQAQPSPMOHQFSM 134
Db 63 YFLRPGDSKKPIIYDVETLRDGNFSARRVAAIQNGKPIFYMTASF-QAPEAGFEHQKTM 121
QY 135 PTVPPELDCETLIDQ---YLRDPNLOKRYPLALNRIAAOEVPIETKPVN-PSPLSQL 190
Db 122 PSAPADG-LPSETQIAQSLAHLPPVLKDKF-----ICDRPLEVRPVEFHNPLKG- 171
QY 191 QRMPEKQMFWRARGYIGEGDMKHCVAAYISDYAFILGTALLPHQWOKVHFH----- 244
Db 172 HVAEPHQVWRANGSVDP-DLRVHQYLLGYASDLNPLVALQP-----HGICFLEPGIOI 226
QY 245 VSLDHSWMFHAPRADHMLYECEPSWAGSGRLVHGLRWQDGLVAVTCAQEGVIR 301
Db 227 ATIDHSMWFRHPFNLEWLLYSVESTSASSARGFVRGFEFTQDGVILVASTVQEGVMR 283

RESULT 6
US-09-265-294-3
; Sequence 3, Application US/09265294
; Patent No. 6210890
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Shah, Purvi
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN PEROXISOMAL THIOESTERASE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; FILING DATE: US/09/265,294
; PRIOR APPLICATION NUMBER:
; APPLICATION NUMBER: 09/100,851
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0293 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 286 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank

CLONE: 147932
US-09-265-294-3

Query Match 33.28; Score 553; DB 4; Length 286;
Best Local Similarity 42.18; Pred. No. 4.2e-55;
Matches 125; Conservative 53; Mismatches 91; Indels 28; Gaps 10;

QY 15 LRSVLVTVLNLEPLDEDLFRGRHYWVPAKRLFGGQIVGOALVAAKSVSDVHVHSLHC 74
Db 5 LKLL--TLNLEKIEGLFRGQSEDLGRVFGGVQVGOALYAARETVPEERLVHSFHS 62
QY 75 YFVRAGDPKLPVLYQVTRTGSFVSRSVKAQVHGKPIFCQAFQQAQPSPMOHQFSM 134
Db 63 YFLRPGDSKKPIIYDVETLRDGNFSARRVAAIQNGKPIFYMTASF-QAPEAGFEHQKTM 121
QY 135 PTVPPELDCETLIDQ---YLRDPNLOKRYPLALNRIAAOEVPIETKPVN-PSPLSQL 190
Db 122 PSAPADG-LPSETQIAQSLAHLPPVLKDKF-----ICDRPLEVRPVEFHNPLKG- 171
QY 191 QRMPEKQMFWRARGYIGEGDMKHCVAAYISDYAFILGTALLPHQWOKVHFH----- 244
Db 172 HVAEPHQVWRANGSVDP-DLRVHQYLLGYASDLNPLVALQP-----HGICFLEPGIOI 226
QY 245 VSLDHSWMFHAPRADHMLYECEPSWAGSGRLVHGLRWQDGLVAVTCAQEGVIR 301
Db 227 ATIDHSMWFRHPFNLEWLLYSVESTSASSARGFVRGFEFTQDGVILVASTVQEGVMR 283

RESULT 7
US-08-872-784-4
; Sequence 4, Application US/08872784
; Patent No. 5776753
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Shah, Purvi
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN PEROXISOMAL THIOESTERASE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/872,784
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0293 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 349 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 854594

US-08-872-784-4

Query Match 16.4%; Score 273; DB 1; Length 349;
Best Local Similarity 26.2%; Pred. No. 6.7e-23;
Matches 86; Conservative 53; Mismatches 135; Indels 54; Gaps 9;

Qy 23 VLNLEPLDEDLFRGRHYWVPA-----KRLFGGQIVGQALVAAAKSVSDVHVHSLHCYFV 77
Db 14 ILELVPLSPTSEVTK--YLPAPVSGKTFGGTLVSQSLLASLHTVPLNFPFPTSLHSYFI 71
Qy 78 RAGDPKPLVLYOVERTRTGSSFSVRSKAVOHGKPIFICOASF--QQAQSPMOHQFSM 134
Db 72 KGGDPRTKITVHVQNLNRGRNFTHKOVSAYQHDKLIFTSMLFAVORSKEHDSLQHWETI 131
Qy 135 PTV-----PPPELLDCETLIDQYLRDNPLOKRYPLALNRIAAQEVPIEIKPVNPSLSQL 190
Db 132 PGLQKQDPHRYEATSLFOKEVLDPOKLSRYASLSDRF--QDATSMKSYVDAFYQVGM 189
Qy 191 QRMPEKQMF-----VRARGYI-----GEGDMK 214
Db 190 EYQFPKDMFYSAARHTDELDFYFKVRPPIITVEHAGDESSLKHHPYRIPKSIITPENDARY 249
Qy 215 HCCVAAVSDYAFGLTALLPHQWQHKVH-FMVSLDHSMMFHFAPFRADHWMLEYCESPWAG 273
Db 250 NYVAFAYLSDSYLLTIPIYFHNLPYCHSFSVSLDHTIIFHQLPHVNNWYILKISNPRSH 309
Qy 274 GSRGLVHGRWL-RDQGVLAATCAQEGVI 300
Db 310 WDKHLVQCKYFDTQSGRIMASVSQEGYV 337

RESULT 8

US-09-100-851-4
; Sequence 4, Application US/09100851
; Patent No. 5911984
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Shah, Purvi
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN PEROXISOMAL THIOESTERASE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,851
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/872,784
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0293 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 349 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: GenBank

CLONE: 854594

US-09-100-851-4

Query Match 16.4%; Score 273; DB 2; Length 349;
Best Local Similarity 26.2%; Pred. No. 6.7e-23;
Matches 86; Conservative 53; Mismatches 135; Indels 54; Gaps 9;

Qy 23 VLNLEPLDEDLFRGRHYWVPA-----KRLFGGQIVGQALVAAAKSVSDVHVHSLHCYFV 77
Db 14 ILELVPLSPTSEVTK--YLPAPVSGKTFGGTLVSQSLLASLHTVPLNFPFPTSLHSYFI 71
Qy 78 RAGDPKPLVLYOVERTRTGSSFSVRSKAVOHGKPIFICOASF--QQAQSPMOHQFSM 134
Db 72 KGGDPRTKITVHVQNLNRGRNFTHKOVSAYQHDKLIFTSMLFAVORSKEHDSLQHWETI 131
Qy 135 PTV-----PPPELLDCETLIDQYLRDNPLOKRYPLALNRIAAQEVPIEIKPVNPSLSQL 190
Db 132 PGLQKQDPHRYEATSLFOKEVLDPOKLSRYASLSDRF--QDATSMKSYVDAFYQVGM 189
Qy 191 QRMPEKQMF-----VRARGYI-----GEGDMK 214
Db 190 EYQFPKDMFYSAARHTDELDFYFKVRPPIITVEHAGDESSLKHHPYRIPKSIITPENDARY 249
Qy 215 HCCVAAVSDYAFGLTALLPHQWQHKVH-FMVSLDHSMMFHFAPFRADHWMLEYCESPWAG 273
Db 250 NYVAFAYLSDSYLLTIPIYFHNLPYCHSFSVSLDHTIIFHQLPHVNNWYILKISNPRSH 309
Qy 274 GSRGLVHGRWL-RDQGVLAATCAQEGVI 300
Db 310 WDKHLVQCKYFDTQSGRIMASVSQEGYV 337

RESULT 9

US-09-265-294-4
; Sequence 4, Application US/09265294
; Patent No. 6210890
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Shah, Purvi
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN PEROXISOMAL THIOESTERASE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/265,294
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,851
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0293 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 349 amino acids

; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 854594
US-09-265-294-4

Query Match 16.4%; Score 273; DB 4; Length 349;
Best Local Similarity 26.2%; Pred. No. 6.7e-23;
Matches 86; Conservative 53; Mismatches 135; Indels 54; Gaps 9;

QY 23 VNLPLEDDEDLPRGRHYWTPA-----KRLFGGQIVGQALVAAAKSVSEDEVHVSLSHCYFV 77
DB 14 ILELPLSPSTSVTK--YLPAPVPGSGTFTGGLVLSQSLASLHTVPLNFPPTSLHSYFI 71
QY 78 RAGDPKPLVLYOVERTRTSSFSVRSKAVQHGKPIFICOASF---QQAQSPSPMOHQFSM 134
DB 72 KGGDPRTKITYHVQNLNRGNRTIHKOVSAIQHDKLIFTSMLFAVQSKSHDSLOHWETI 131
QY 135 PTV-----PPPELLDCETLIDQYLRDPNLQKRYPLALNRILAAQEVPIEIKPVNPSPLSQL 190
DB 132 PGLQCKQDPHRYEATSLFQREVLPQKLSRYASLSDRF--QDATSMKYVDFAFOYGV 189
QY 191 QRMPEKQFW-----VRARGYI-----GEGDMKM 214
DB 190 EYQFPKDMFYSAHRTDELDFYFKVRPPITTVHAGDESSLHKHHPYRIPKSIPTENDARY 249
QY 215 HCCVAAIYSDYAFGLTALLPHQWQKHVH-FMVSLSHSMFHPFRADHWMLYECESPWAG 273
DB 250 NYVAFAYLSDSVLLTIPYFHNLPYCHFSVSLDHTIYFPHQPLPHVNNWNYLKLISNPRSH 309
QY 274 GSRGLVHGLRW-RQDGLVAVTCAQEGVI 300
DB 310 WDKHLVQCKYEDTQSGRIMASVSQEGYV 337

RESULT 10
US-08-804-227C-11

; Sequence 11, Application US/08804227C
; Patent No. 5876991
; GENERAL INFORMATION:
; APPLICANT: DeHoff, Bradley S.
; APPLICANT: Kuhnstoss, Stuart A.
; APPLICANT: Rostock, Paul R., Jr.
; APPLICANT: Sutton, Kimberly L.
; TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: THOMAS G. PLANT 1501
; STREET: LILLY CORPORATE CENTER
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: USA
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: ASCII(DOS) Text only
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/804.227C
; FILING DATE: February 21, 1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: plant, Thomas, G.
; REGISTRATION NUMBER: 35,784
; REFERENCE/DOCKET NUMBER: X-8231
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-2459
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:

; LENGTH: 1580 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-804-227C-11

Query Match 6.0%; Score 100; DB 2; Length 1580;
Best Local Similarity 20.6%; Pred. No. 0.042;
Matches 66; Conservative 25; Mismatches 122; Indels 108; Gaps 11;

QY 28 PLDEDLFRGRHYWVPAKRLFGGQIVGQALVAAAKSVSEDEVHVSLSHCYFVRAGDPKLPVL 87
DB 900 PLPTAFASDRDYWLPT-----AIGRRVAAAPV-----DASAPGR 934
QY 88 YOVERTRTSGSFSVRSKAVQHGKPIFICOASFQQAQSPSPMOHQFSMPTVPPE----- 141
DB 935 YRVWT-----PV-----ASDSSGRPSGRWLLVQTFGTAPDEADTAAS 972
QY 142 ELLDCETLIDQYLRDPNLQKRYPLA--LNRIAAQ-----EYPIEIKPVN 183
DB 973 ALGAAGVVERCLLDPTAAARVTLTERLAELDAQPEGLAGVLVLPGRPOSTAPADASPLD 1032
QY 184 PPSLSLOQME-----PKMFVVRARGYIGEGDKMHCCVAAIYSDYAFGLTALLPHQW 237
DB 1033 PCTAAVLLVYQVAPDAAPKARIWYTRGAVAGSGEYPCAYGARVMGLRVAALVPPVQM 1092
QY 238 QHKVHFVMSLDHSMFHPFRADHWMLYECESPWAGSR-----GLVHGR----- 282
DB 1093 GGLVDVAVGAGVREW-----RRVGVVAGGEDQVAVRGGVFRRLVGVGR 1140
QY 283 -----LWRQDGLVAVTCAQEGV 299
DB 1141 GSGVVRARGCVVVTGGLGV 1161

RESULT 11

US-08-804-198-5
; Sequence 5, Application US/08804198
; Patent No. 5945320
; GENERAL INFORMATION:

; APPLICANT: Burgett, Stanley G.
; APPLICANT: Kuhnstoss, Stuart A.
; APPLICANT: Rao, Nagaraja R.
; APPLICANT: Richardson, Mark A.
; APPLICANT: Rostock, Paul R., Jr.
; TITLE OF INVENTION: PLATENOLIDE SYNTHASE GENE
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PAUL R. CANTRELL 1138
; STREET: LILLY CORPORATE CENTER
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: USA
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: Macintosh 7.0
; SOFTWARE: Microsoft Word 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/804.198
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: CANTRELL, PAUL R.
; REGISTRATION NUMBER: 36,470
; REFERENCE/DOCKET NUMBER: P9113
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-3885
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1580 amino acids

```

; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-804-198-5

Query Match      6.0%; Score 100; DB 2; Length 1580;
Best Local Similarity 20.6%; Pred. No. 0.042;
Matches 66; Conservative 25; Mismatches 122; Indels 108; Gaps 11;

QY 28 PLDEDLFRGRHWYPAKRLFGQIYGOALVAAAKSVSEDDVHVSLSHCYFVRAGDPKLPVL 87
DB 900 PLPTVAFSRDRYWLPT-----AIGRAVEAAPV-----DASAPGR 934

QY 88 YQVETFRGSSFSVRSKAVQHGKPIFCOASFQAQSPMQHQSMPVTPPPE----- 141
DB 935 YRVVTWT-----PV-----ASDGRSPSGRWLLVOTPGTAPDEADTAAS 972

QY 142 ELLDCETLIDYLRPNLQKRYPLA--LNRIAAQ-----EVPIETKPVN 183
DB 973 ALGAAGVVVERCLLOPTEARVTLTERLAELDAQPEGLAGVLVLPGRPOSTAPADASPLD 1032

QY 184 PPSLSQLRME-----PKQFWVRARGVIGEDMKMHCCVAAVYSDYAFGLGTALLPHQW 237
DB 1033 PGTAALLVQVAVPDAAPKARIWVTVTRGAVAVGSEVPCAVGARVWGLGRVAALEVPVQW 1092

QY 238 QHKVHFVMSLDHSMFHAFPRADHMLYECESPWAGGSR-----GLVHGR----- 282
DB 1093 GGLVDVAVGAGVREW-----RRVGVVAGGEDQVAVRGGVGRRLLVGVGVR 1140

QY 283 -----LWRQDQVLAVTCQAEV 299
DB 1141 GSGGVWRARGCVVVTGGLGV 1161

RESULT 12
US-08-417-089-6
; Sequence 6, Application US/08417089
; Patent No. 6069298
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: METHODS AND AN ACETYL COA CARBOXYLASE GENE
; TITLE OF INVENTION: FOR CONFERRING HERBICIDE TOLERANCE AND AN ALTERATION IN
; TITLE OF INVENTION: OIL CONTENT OF PLANTS
; NUMBER OF SEQUENCES: 11
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION NUMBER: US/08/417,089
; APPLICATION FOR SEQ ID NO: 6:
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2325 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-417-089-6

Query Match      5.3%; Score 89; DB 3; Length 2325;
Best Local Similarity 19.9%; Pred. No. 1.4;
Matches 57; Conservative 43; Mismatches 102; Indels 84; Gaps 13;

QY 26 LEPLDEDLFRGRHWYPAKRLFGQIYGOALVAAAKSVSEDDVHVSLSHCYFVRAGDPKLP 85
DB 287 LDAIPEEMRKACVTTTEEAASVAVSGVYGPAMIKASMGSGGKGRKYVH-----NDDEV 340

QY 86 VLQVETRTGSSFSVRSKAVQHGKPIFCOASFQAQSPMQHQSMPVTPPPEELLDD 145
DB 341 ALF-----KQVQGVGSPGFVNRILASQS-----RH-----LE 368
```

```

QY 146 CETLIDY-----LRDPNLQKRYPLALNRIAAQEVPIETKPVNPSPLSQLRMPEKQM 198
DB 369 VOLLCDQYGNVAALHSRDCSVQRH-----QKI-IEEGPVTVPAPRETVKALEQAAR 418

QY 199 FWVRARGYIGEGDMKMHCCVAAVYSDYAFGLGTALLPH-QWQHKV-----HFV 245
DB 419 RLAKAGVYVGAARVEY--LYSMETGDIYEL--ELNPLQVHEHPVTEWIAEVNLPAAQVAV 474

QY 246 SLDHSMFHAFPRADHMLYECESPWAGGSRGLVHGRRLWRQDGVLA 291
DB 475 GMGIPLWQIPEIRRFYGMDF-----GGGYD-----IWRKTAALA 508

RESULT 13
US-08-695-651-6
; Sequence 6, Application US/08695651
; Patent No. 6146867
; GENERAL INFORMATION:
; APPLICANT: Gengenbach, B. G.
; APPLICANT: Somers, D. A.
; APPLICANT: Wyse, D. L.
; APPLICANT: Gronwald, J. W.
; APPLICANT: Egli, M. A.
; APPLICANT: Lutz, S. M.
; TITLE OF INVENTION: METHOD FOR PRODUCING PLANT ACETYL COA
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.
; STREET: P.O. Box 2938
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION NUMBER: US/08/695,651
; FILING DATE: 12-AUG-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/417089
; FILING DATE: 05-APR-1995
; APPLICATION NUMBER: 08/014326
; FILING DATE: 05-FEB-1993
; APPLICATION NUMBER: 07/917462
; FILING DATE: 21-JUL-1992
; APPLICATION NUMBER: 07/538674
; FILING DATE: 18-JUN-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Woessner, Warren D
; REGISTRATION NUMBER: 30,440
; REFERENCE/DOCKET NUMBER: 600.318US3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-373-6900
; TELEFAX: 612-339-3061
; TELEX:
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2325 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: N-terminal
US-08-695-651-6

Query Match      5.3%; Score 89; DB 4; Length 2325;
Best Local Similarity 19.9%; Pred. No. 1.4;
```

	Matches	57; Conservative	43; Mismatches	102; Indels	84; Gaps	13;
QY	26	LEPLEDEDLFRGRHYWPAKRLFGQIQVGOALVAARAAKSVSEDEVHVRSLRUCYFVRAGDPKLP	85			
Db	287	LDAIPEMYRKACVTTTEEAASQCVVGYGPAMIKASMGGGKGLRKVH-----NDDEV	340			
QY	86	VLYQVETRTRTSSFSYRSYKAVGHGKPIFCIQASFOQAQPSMQHQSFMTPVPPPELLD	145			
Db	341	ALF-----KQVGEVPGSPFVNRUASO-----RH-----LE	368			
QY	146	CEFLIDQY-----LRDPLNLRKYPLALNRIRAAQEVPIEIKVPNSPLSQLQRMPEKQM	198			
Db	369	VOILCDQYGNVAALHSRDCDSVORRH-----QKI-IEEGPVTVAPRETVKALEQAAR	418			
QY	199	FWVRARGYIGEGDMKMHCCVAAVYSYAFLLGTALLPH-QWQHKV-----HFV	245			
Db	419	RLAKAVGYVGAATVEY--LYSMETGDYFL--ELNPLQVEHPTWTWIAEVNPLPAAQVAV	474			
QY	246	SLDHSWFIAPFRADHIMWLYECESPWAGSGRLVGHGLRMQRDGVLA	291			
Db	475	GMGIPLWQIPEIRRFYGMDY-----GGYD-----IWRKTAALA	508			

```

RESULT 14
US-08-930-285-6
; Sequence 6, Application US/08930285
; Patent No. 6222099
; GENERAL INFORMATION:
; APPLICANT: Regents of the University of Minnesota, et al.
; TITLE OF INVENTION: TRANSGENIC PLANTS EXPRESSING ACETYL COA CARBOXYLASE GE
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P. A.
; STREET: P. O. Box 2938
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/930,285
; FILING DATE: 13-APR-1998
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/04625
; FILING DATE: 04-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Embretson, Janet E.
; REGISTRATION NUMBER: 39,665
; REFERENCE/DOCKET NUMBER: 600.318US4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-339-0331
; TELEFAX: 612-339-3061
; TELEX:
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2325 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEetical: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal and C-terminal (full length protein)
; ORIGINAL SOURCE:
US-08-930-285-6

```

Query Match 5.3%; Score 89; DB 4; Length 2325;

```

Best Local Similarity 19.9%; Pred. No. 1.4;
Matches 57; Conservative 43; Mismatches 102; Indels 84; Gaps 13;

Qy 26 VLQEDLPRGRHYWPAKFLFGQIVTQGOALVAAKSVSDVHVHSLHCYFVRGDPKLP 85
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 287 LDAIPEMYRKACVTTTEEAVASQVVGYPAMIKASGGGGKGIRKVVH-----NDDEV 340

Qy 86 VLYQVETRTGTSFSYRSVKAVOGRKPIFCIQASFOQAQSPMQHQFSNPTVPPDEELD 145
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 341 ALF-----KQVQGEVPGSPFVWRLASQS-----RH-----LE 368

Qy 146 CETLIDQY-----LRDPNLQRYPLALNRIAAQVEPIEKPNPSPLSQLORNEPKQM 198
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 369 VOLLCDQYGNVALHSRDCSVQRH-----QKI-IEEGPVTAPRETVKALEQAR 418

Qy 199 FWVRARGYIGEDMKHKCCVAAVYISDYAFGLTALLPH-QWQHKV-----HFW 245
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 419 RLAKAGVYGAAATVEY--LYSNETGDYFL--ELNPRLOVEHPVTEIAEYNLPAQAV 474

Qy 246 SLDSHMWFHAPRADHMVLYECESPWAGSGRGLVHGRLWRQDGVLA 291
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 475 GNGIPLWQIPEIRRFYGMGY-----GGGYD-----IWRKTAALA 508

```

RESULT 15
US-08-695-421-6
: Sequence 6, Application US/08695421
: Patent No. 6268550
: GENERAL INFORMATION:
: APPLICANT: Gengenbach, B. G.
: APPLICANT: Somers, D. A.
: APPLICANT: Wyse, D. L.
: APPLICANT: Gronwald, J. W.
: APPLICANT: Egli, M. A.
: APPLICANT: Lutz, S. M.
: TITLE OF INVENTION: METHODS AND AN ACETYL COA CARBOXYLASE GENE
: NUMBER OF SEQUENCES: 11
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P. A.
: STREET: P. O. Box 2938
: CITY: Minneapolis
: STATE: MN
: COUNTRY: USA
: ZIP: 55402
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSEQ Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/695,421
: FILING DATE: 23-AUG-1996
: CLASSIFICATION: 800
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/417089
: FILING DATE: 05-APR-1995
: APPLICATION NUMBER: 08/014326
: FILING DATE: 05-FEB-1993
: APPLICATION NUMBER: 07/917462
: FILING DATE: 21-JUL-1992
: APPLICATION NUMBER: 07/538674
: FILING DATE: 18-JUN-1990
: ATTORNEY/AGENT INFORMATION:
: NAME: Woessner, Warren D
: REGISTRATION NUMBER: 30,440
: REFERENCE/DOCKET NUMBER: 600.318052
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 612-373-6900
: TELEFAX: 612-339-5061
: TELEX:
: INFORMATION FOR SEQ ID NO: 6:
: SEQUENCE CHARACTERISTICS:

